ANALYSIS & PDE

Volume 8 No. 6 2015

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2015 is US $205/year for the electronic version, and $390/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers 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. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1289

msp

ON SMALL ENERGY STABILIZATION IN THE NLS WITH A TRAPPING POTENTIAL

SCIPIO CUCCAGNA AND MASAYA MAEDA

We describe the asymptotic behavior of small energy solutions of an NLS with a trapping potential, generalizing work of Soffer and Weinstein, and of Tsai and Yau. The novelty is that we allow generic spectra associated to the potential. This is a new application of the idea of interpreting the *nonlinear Fermi golden rule* as a consequence of the Hamiltonian structure.

1. Introduction

We consider the initial value problem

$$\mathrm{i}u_t = Hu + |u|^2 u, \quad (t,x) \in \mathbb{R}^{1+3}, \qquad u(0) = u_0, \tag{1-1}$$

where $H = -\Delta + V$. For $f, g : \mathbb{R}^3 \to \mathbb{C}$, we introduce the bilinear form

$$\langle f, g\rangle = \int_{\mathbb{R}^3} f(x)g(x)\,dx. \tag{1-2}$$

We assume the following:

(H1) $V \in \mathcal{S}(\mathbb{R}^3)$, where $\mathcal{S}(\mathbb{R}^3)$ is the space of Schwartz functions.

(H2) $\sigma_p(H) = \{e_1 < e_2 < e_3 < \cdots < e_n < 0\}$. Here we assume that all the eigenvalues have multiplicity 1. Zero is neither an eigenvalue nor a resonance (that is, if $(-\Delta + V)u = 0$ with $u \in C^\infty$ and $|u(x)| \le C|x|^{-1}$ for a fixed C, then $u = 0$).

MSC2010: 35Q55.
Keywords: nonlinear Schroedinger equation, asymptotic stability.

(H3) There is an $N \in \mathbb{N}$ with $N > |e_1|(\min\{e_i - e_j : i > j\})^{-1}$ such that, if $\mu \in \mathbb{Z}^n$ satisfies $|\mu| \le 4N+8$ and $\boldsymbol{e} := (e_1, \dots, e_n)$, then we have

$$\mu \cdot \boldsymbol{e} := \mu_1 e_1 + \cdots + \mu_n e_n = 0 \iff \mu = 0 .$$

(H4) The following Fermi golden rule (FGR) holds: the expression

$$\sum_{L \in \Lambda} \langle \delta(H-L)\overline{G}_L(\zeta), G_L(\zeta)\rangle,$$

which is defined in the course of the paper (for $\Lambda \subset \mathbb{R}_+$ see (6-25) and for G_L see (6-44)) and which is always nonnegative, satisfies formula (6-47).

To each e_j we associate an eigenfunction ϕ_j. We choose them so that $\langle \phi_j, \overline{\phi}_k\rangle = \delta_{jk}$ and, since we can, we also choose the ϕ_j to be all real valued. To each ϕ_j we associate nonlinear bound states.

Proposition 1.1 (bound states). *Fix $j \in \{1, \dots, n\}$. Then there exists $a_0 > 0$ such that, for all $z \in B_{\mathbb{C}}(0, a_0)$, there is a unique $Q_{jz} \in \mathcal{S}(\mathbb{R}^3, \mathbb{C}) := \bigcap_{t \ge 0} \Sigma_t(\mathbb{R}^3, \mathbb{C})$ (for the spaces Σ_t, see Section 2) such that*

$$HQ_{jz} + |Q_{jz}|^2 Q_{jz} = E_{jz} Q_{jz}, \quad Q_{jz} = z\phi_j + q_{jz}, \quad \langle q_{jz}, \overline{\phi}_j\rangle = 0, \tag{1-3}$$

and such that we have, for any $r \in \mathbb{N}$:

(1) *$(q_{jz}, E_{jz}) \in C^\infty(B_{\mathbb{C}}(0, a_0), \Sigma_r \times \mathbb{R})$, $q_{jz} = z\hat{q}_j(|z|^2)$ with $\hat{q}_j(t^2) = t^2 \tilde{q}_j(t^2)$, where $\tilde{q}_j(t)$ is in $C^\infty((-a_0{}^2, a_0{}^2), \Sigma_r(\mathbb{R}^3, \mathbb{R}))$, and $E_{jz} = E_j(|z|^2)$ with $E_j(t) \in C^\infty((-a_0{}^2, a_0{}^2), \mathbb{R})$.*

(2) *$\|q_{jz}\|_{\Sigma_r} \le C|z|^3$, $|E_{jz} - e_j| < C|z|^2$ for some $C > 0$.*

For the proof of Proposition 1.1 see Appendix A.

Definition 1.2. Let $b_0 > 0$ be sufficiently small so that, for $z_j \in B_{\mathbb{C}}(0, b_0)$, the function Q_{jz_j} exists for all $j \in \{1, \dots, n\}$. For such z_j and for D_{jI} and D_{jR}, defined in Section 2, we set

$$\mathcal{H}_c[z] = \mathcal{H}_c[z_1, \dots, z_n] := \bigl\{\eta \in L^2 : \mathrm{Re}\langle \mathrm{i}\overline{\eta}, D_{jR} Q_{jz_j}\rangle = \mathrm{Re}\langle \mathrm{i}\overline{\eta}, D_{jI} Q_{jz_j}\rangle = 0 \text{ for all } j\bigr\}. \tag{1-4}$$

In particular, as an elementary consequence of (1-4) and Proposition 1.1, we have

$$\mathcal{H}_c[0] = \{\eta \in L^2 : \langle \overline{\eta}, \phi_j\rangle = 0 \text{ for all } j\}. \tag{1-5}$$

We denote by P_c the orthogonal projection of L^2 onto $\mathcal{H}_c[0]$.

A pair (p, q) is *admissible* when

$$\frac{2}{p} + \frac{3}{q} = \frac{3}{2}, \quad 6 \ge q \ge 2, \quad p \ge 2. \tag{1-6}$$

The following theorem is our main result:

Theorem 1.3. *Assume* (H1)–(H4). *Then there exist $\epsilon_0 > 0$ and $C > 0$ such that, if $\epsilon = \|u(0)\|_{H^1} < \epsilon_0$, the solution $u(t)$ of (1-1) can be written uniquely for all times as*

$$u(t) = \sum_{j=1}^{n} Q_{jz_j(t)} + \eta(t) \quad \textit{with} \quad \eta(t) \in \mathcal{H}_c[z(t)] \tag{1-7}$$

in such a way that there exist a unique j_0, a $\rho_+ \in [0,\infty)^n$ with $\rho_{+j} = 0$ for $j \neq j_0$ and $|\rho_+| \le C\|u(0)\|_{H^1}$, and an $\eta_+ \in H^1$ with $\|\eta_+\|_{H^1} \le C\|u(0)\|_{H^1}$, such that

$$\lim_{t\to+\infty} \|\eta(t,x) - e^{\mathrm{i}t\Delta}\eta_+(x)\|_{H^1_x} = 0, \qquad \lim_{t\to+\infty} |z_j(t)| = \rho_{+j}. \tag{1-8}$$

Furthermore, we have $\eta = \tilde{\eta} + A(t,x)$ such that, for all admissible pairs (p,q),

$$\|z\|_{L^\infty_t(\mathbb{R}_+)} + \|\tilde{\eta}\|_{L^p_t(\mathbb{R}_+, W^{1,q}_x)} \le C\|u(0)\|_{H^1} \quad \textit{and} \quad \|\dot{z}_j + \mathrm{i}e_j z_j\|_{L^\infty_t(\mathbb{R}_+)} \le C\|u(0)\|^2_{H^1} \tag{1-9}$$

and such that $A(t,\cdot) \in \Sigma_2$ for all $t \ge 0$ and

$$\lim_{t\to+\infty} \|A(t,\cdot)\|_{\Sigma_2} = 0. \tag{1-10}$$

As an interesting corollary to Theorem 1.3, we show rather simply that the excited states are *orbitally unstable*. We recall that $e^{-\mathrm{i}tE_{jz}}Q_{jz}$ is called *orbitally stable* in $H^1(\mathbb{R}^3)$ for (1-1) if

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \quad \|u_0 - Q_{jz}\|_{H^1(\mathbb{R}^3)} < \delta \implies \sup_{t\in\mathbb{R}} \inf_{\vartheta\in\mathbb{R}} \|u(t) - e^{\mathrm{i}\vartheta}e^{-\mathrm{i}tE_{jz}}Q_{jz}\|_{H^1(\mathbb{R}^3)} < \varepsilon \tag{1-11}$$

and is orbitally unstable if (1-11) does not hold. We prove:

Theorem 1.4. *Assume* (H1)–(H4). *Then there exists $\epsilon_0 > 0$ such that, if $j \ge 2$, and for $|z| < \epsilon_0$, the standing wave $e^{-\mathrm{i}tE_{jz}}Q_{jz}$ is orbitally unstable. Furthermore, $e^{-\mathrm{i}tE_{1z}}Q_{1z}$ is orbitally stable.*

Notice that [Tsai and Yau 2002b; 2002c; 2002d; Soffer and Weinstein 2004; Gang and Weinstein 2008; 2011; Gustafson and Phan 2011; Nakanishi et al. 2012] contain only very partial proofs of the instability of the second excited state. Theorem 1.4 will be proved in Section 7 and, until then, and in particular in the sequel of this introduction, we will focus only on Theorem 1.3.

We recall that [Gustafson et al. 2004] proved Theorem 1.3 for $|u|^2u$ replaced by more general functions in the case when H has one eigenvalue (for the NLS with an electromagnetic potential, we refer to [Koo 2011]). The case of two eigenvalues is discussed in the series [Tsai and Yau 2002a; 2002b; 2002c] and in [Soffer and Weinstein 2004] under more stringent conditions on the initial data, which are such that $\|u_0\|_{H^{k,s}}$ is small for $k > 2$ and some s large enough in [Soffer and Weinstein 2004] and $\|u_0\|_{H^1\cap L^{2,s}}$ small for $s > 3$ in [Tsai and Yau 2002a; 2002b; 2002c]. A crucial restriction in these papers is that $2e_2 > e_1$. They then prove versions of Theorem 1.3 involving also rates of decay of $|z(t)|$, of $\|\eta(t)\|_{L^\infty(\mathbb{R}^3)}$, and of $\|\eta(t)\|_{L^{2,s}(\mathbb{R}^3)}$ for appropriate $s > 0$.

The ideas used in proofs such as in [Tsai and Yau 2002a; 2002b; 2002c; Soffer and Weinstein 2004] appear to be very difficult to extend to operators with more than 2 eigenvalues, where only partial results like in [Nakanishi et al. 2012] are known, and for initial data small only in H^1. On one hand, the Poincaré–Dulac normal form argument in these papers seems not suited to discuss the higher-order FGR needed when $2e_2 < e_1$. Furthermore, in these papers there is a subdivision of the evolution into distinct phases, which the solution enters in a somewhat irreversible fashion and which are considered one by one. This division into distinct phases might become unclear in cases when $u(t)$ oscillates from one phase to the other, as is not unlikely to happen in the H^1 case, or when the passage from one phase to the other is very slow, as is certainly true in the H^1 case. Moreover, an increase in the number of eigenvalues of H increases also the number of distinct phases that need to be accounted for and the complexity of the argument. So, any hope of proving Theorem 1.3 should rely on an argument which yields the asymptotics in a single stroke and which does not distinguish distinct cases. This is what we do; see, for example, the second part of Section 6. We did not check if our method yields the decay estimates of [Tsai and Yau 2002a; 2002b; 2002c; Soffer and Weinstein 2004] under more stringent conditions on u_0.

We give a new application of the interpretation of the FGR in terms of the Hamiltonian structure of the equation. This interpretation was first introduced in [Cuccagna 2009] and was then applied in [Bambusi and Cuccagna 2011] to generalize the result of [Soffer and Weinstein 1999]. It was later applied to the problem of asymptotic stability of ground states of the NLS, first not allowing translation symmetries in [Cuccagna 2011a], and then with translation in [Cuccagna 2014]; see also [Cuccagna 2012].

The link between FGR and Hamiltonian structure rests in the fact that the latter yields algebraic identities between coefficients of different coordinates in the system (compare the right-hand side in (6-13) with the second line in (6-27)). These allow us to show that some other coefficients in the equations of the z_j have a square power structure and have a fixed sign (in the case of the NLS); see Lemma 6.8. This then yields decay of the z_j, except for at most one of the j here. We refer to pp. 287–288 in [Cuccagna 2011a] for the original intuition behind this approach to the FGR, which views the FGR as a simple consequence of Schwartz's lemma on mixed derivatives, and which has been used in [Bambusi and Cuccagna 2011; Cuccagna 2009; 2011a; 2014; 2012], among others. For other applications of this theory we refer to the references in [Cuccagna 2012; Cuccagna and Maeda 2014]. We refer also to [Cuccagna 2011b], whose treatment of the FGR is similar to the one in this paper. Earlier treatments of FGR, are in [Tsai and Yau 2002a; 2002b; 2002c; Soffer and Weinstein 2004] and, still earlier, in [Buslaev and Perel′man 1995; Soffer and Weinstein 1999], but they seem to work only in relatively simple cases, because they run into trouble if the normal form argument requires more than a very few steps. For more references and comments see [Cuccagna 2011a].

As we will see below, the FGR can be seen relatively easily after one finds an appropriate effective Hamiltonian in the right system of coordinates. This coordinate system is obtained by a normal form argument. Right from the beginning, though, it is crucial to choose the right ansatz and system of coordinates. For example, since H has eigenvalues, it would seem natural to split the NLS (1-1) into a system using the coordinates of the spectral decomposition of H; see (4-2). However, this would not be a good choice for our nonlinear system. Following [Gustafson et al. 2004], it is better to pick as coordinates the z_j of Proposition 1.1, complementing them with an appropriate continuous coordinate. There is the natural ansatz (2-1) (the same used in [Soffer and Weinstein 2004]), which, following [Gustafson et al. 2004], can be used to obtain the continuous coordinate, here denoted η and introduced in Lemma 2.4.

Once we have coordinates (z, η) with $z = (z_1, \dots, z_n)$, where z_1 is the ground state coordinate, z_j for $j > 1$ the excited states coordinates, and η the radiation coordinate, Theorem 1.3 can be loosely paraphrased as

$$\eta(t) \to 0 \quad \text{in } H^1_{\mathrm{loc}} \quad \text{and} \quad z_j(t) \to 0 \quad \text{except for at most one } j. \tag{1-12}$$

In particular, if $z(t) \to 0$ the solution $u(t)$ of (1-1) scatters like a solution of $\mathrm{i}\dot{u} = -\Delta u$ in H^1. Otherwise there is one j such that $u(t)$ scatters to $e^{\mathrm{i}\vartheta(t)} Q_{z_{+j}}$, with $\vartheta(t)$ a phase term which we do not control here. We have convergence by scattering to a ground state if $j = 1$, and to an excited state if $j > 1$. The latter presumably occurs for the $u(t)$ whose trajectory is contained in an appropriate manifold; see [Tsai and Yau 2002d; Beceanu 2012; Gustafson and Phan 2011].

It is not easy to see (1-12) in the initial coordinate system. So we need a Birkhoff normal form argument to identify an effective Hamiltonian, like in [Bambusi and Cuccagna 2011]. Unlike there, but like in [Cuccagna 2011a], the initial coordinates, while quite natural from the point of view of the NLS (1-1), are not Darboux coordinates for the natural symplectic form Ω in the problem; see (4-1). Hence, before doing normal forms, we have first to implement the Darboux theorem to diagonalize the problem (of course, the coordinates arising from the spectral decomposition of H — see (4-2) — are Darboux coordinates, but, as we wrote, they are not suited for our nonlinear asymptotic analysis). So in this paper we need to perform a number of coordinate changes: first a Darboux theorem and then normal form analysis. At the end of the process we get new coordinates $(z_1, \dots, z_n, \eta)$ where the Hamiltonian is sufficiently simple that we can prove (1-12) relatively easily using the FGR (which tells us that all the z_j, except at most one, are damped) and a semilinear NLS for η that shows scattering of η because of linear dispersion. In the context of the theory developed in [Bambusi and Cuccagna 2011; Cuccagna 2011a] and other literature, the work in the last system of coordinates, that is, all the material in Section 6, is rather routine.

Having proved (1-12) for the last system of coordinates (z, η), the obvious question is why (1-12) should hold, as Theorem 1.3 is saying, also for the initial coordinates, which we now denote by (z', η') to distinguish them from the final coordinates (z, η). Keeping in mind that all coordinate changes are small nonlinear perturbations of the identity, the only simple reason why this might happen is that different coordinates must be related in the form

$$\begin{aligned} z'_k &= z_k + O(z\eta) + O(\eta^2) + \sum_{i \neq j} O(z_i z_j) \quad \text{for } k = 1, \dots, n, \\ \eta' &= \eta + O(z\eta) + O(\eta^2) + \sum_{i \neq j} O(z_i z_j). \end{aligned} \tag{1-13}$$

This relation between any two systems of coordinates forbids relations like $z'_1 = z_1 + z_2^2$. Indeed, with the latter relations it would not be true (except for the case $z(t) \to 0$) that (1-12) for (z, η) implies (1-12) for (z', η'). So our main strategy is to prove (1-12) for the final (z, η) with some relatively standard method using FGR and linear dispersion, and to be careful to implement only coordinate changes like in (1-13). This latter point is the novel problem we need to face in this paper. It is not obvious from the outset that (1-13) should hold.

As we wrote above, [Gustafson et al. 2004] suggests a very natural choice of functions z_j, based on Proposition 1.1, which can be completed in a system of independent coordinates. Loosely speaking, the z_j have the problem that they are defined somewhat independently to each other. This shows up in the expansion of the Hamiltonian in Lemma 3.1, with a certain lack of decoupling inside the energy between distinct z_j; see (3-9) and Remark 3.2. This leads in (3-3) (see the second line) to terms whose elimination in a normal form argument would seem incompatible with coordinate changes satisfying (1-13). These bad terms of the energy can be better seen in (4-45): they are the $l = 0$ terms in the third line. Other additional bad terms arise in the course of the Darboux theorem transformation. Bad terms in the differential form Γ in (4-17) (used in the classical formula (4-40)) are those in I_1 in (4-22). Specifically, they are the first term in the right-hand side of (4-22). The right-hand side of (4-28) is also filled with bad terms, in the sense that they yield a coordinate change $\mathfrak{F}$ in Lemma 4.8 leading to more $l = 0$ terms in the third line in (4-45). Specifically, they originate from the pullback $\mathfrak{F}^* \sum_{j=1}^{n} E(Q_{jz_j})$ of the first term in the right-hand side of (3-3) (more bad terms seem to arise if we use Ω_0' — see (4-8) — rather than the slightly more complicated Ω_0 — see (4-13) — as the local model of Ω). In a somewhat empirical fashion, for which we don't have a simple conceptual reason, a plain and simple computation shows that all the bad terms cancel out and that there are no $l = 0$ terms in (4-45). This is proved in the cancellation lemma, Lemma 4.11, which is the main new ingredient in the paper. This lemma proves that the change of coordinates designed to diagonalize Ω is also decoupling the discrete coordinates inside the Hamiltonian. From that point on, the structure (1-13) for the coordinate changes is automatic and the various steps of the proof of Theorem 1.3 are similar to arguments such as [Cuccagna 2011b; 2012], which have been repeated in a number of papers. So they are fairly standard, even though we are able to discuss them only in a rather technical way. We have to go into the details of the proof, rather than refer to the references, because of some technical novelties required by the fact that in general $z \nrightarrow 0$, and what converges to 0 is instead the vector $\boldsymbol{Z}$ introduced in Definition 2.2, whose components are products of distinct components of z.

In the second part of Section 6, the FGR and the asymptotics of the z_j in the final coordinate system are rather simple to see in a single stroke. Furthermore, Theorem 6.1 is more or less the same as [Cuccagna 2011a; 2011b].

One limitation in our present paper is that we do not generate examples of equations which satisfy hypothesis (H4). Notice though that our result, for solutions only in H^1, is new even in the 2-eigenvalues case of [Tsai and Yau 2002a; 2002b; 2002c; Soffer and Weinstein 2004], where our FGR is the same. Still, we believe that (H4) holds for generic V. And even if it fails at one stage, this is not necessarily a problem: the strict positive sign in the FGR is only an obstruction to performing further the normal form argument, so, if there is a 0, in principle it is enough to proceed with some further coordinate change until, after a finite number of steps, there will finally be a positive sign in the FGR, and so the stabilization will occur, just at a slower rate. And if the FGR is always 0, then maybe this is because the NLS has a special structure; see [Soffer and Weinstein 1999, p. 69] for some thoughts.

Proposition 2.2 of [Bambusi and Cuccagna 2011] proves validity in general of the FGR. Transposing here that proof would require replacing the cubic nonlinearity with a more general nonlinearity $\beta(|u|^2)u$. This seems rather simple to do because the cubic power is only used to simplify the discussion in Lemma 3.1. But it is not so clear how to offset here the absence of a meaningful mass term m^2u, which in [Bambusi and Cuccagna 2011, pp. 1444–1445], by choosing m generic, is used to move some appropriate spheres in phase space. Adding to the NLS a term m^2u would not change the spheres here.

We reiterate that Proposition 1.1 is valid for small $z_j \in \mathbb{C}$. As z_j increases there are interesting symmetry-breaking bifurcation phenomena; see [Kirr et al. 2008; 2011] and references therein and see also [Fukuizumi and Sacchetti 2011; Grecchi et al. 2002; Sacchetti 2005] and references therein for the semiclassical NLS. Notice that Theorem 1.3 should allow one to prove asymptotic breakdown of the beating motion in the case $\mu_\infty = 0$ in [Grecchi et al. 2002]. Finite-dimensional approximations of the solutions at energies close to the symmetry breaking point of [Kirr et al. 2008] have been considered by [Goodman 2011; Marzuola and Weinstein 2010], who prove the long time existence of interesting patterns for the full NLS. Unfortunately, it is beyond the scope of our analysis, and it remains an interesting open problem, to understand the eventual asymptotic behavior of the solutions in [Goodman 2011; Marzuola and Weinstein 2010].

2. Notation, coordinates and resonant sets

Notation.

- We denote by $\mathbb{N} = \{1, 2, \dots\}$ the set of natural numbers and set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
- We denote $z = (z_1, \dots, z_n)$, $|z| := \sqrt{\sum_{j=1}^n |z_j|^2}$.
- Given a Banach space X, $v \in X$ and $\delta > 0$, we set $B_X(v, \delta) := \{x \in X : \|v - x\|_X < \delta\}$.
- Let A be an operator on $L^2(\mathbb{R}^3)$. Then $\sigma_p(A) \subset \mathbb{C}$ is the set of eigenvalues of A and $\sigma_e(A) \subset \mathbb{C}$ is the essential spectrum of A.
- For $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$, we denote by $\Sigma_r = \Sigma_r(\mathbb{R}^3, \mathbb{K})$ for $r \in \mathbb{N}_0$ the Banach spaces defined by the completion of $C_c(\mathbb{R}^3, \mathbb{K})$ by the norms
$$\|u\|_{\Sigma_r}^2 := \sum_{|\alpha| \le r} (\|x^\alpha u\|_{L^2(\mathbb{R}^3)}^2 + \|\partial_x^\alpha u\|_{L^2(\mathbb{R}^3, \mathbb{K})}^2).$$
For $m < 0$ we consider the topological dual $\Sigma_m = (\Sigma_{-m})'$. Notice — see [Cuccagna 2014] — that the spaces Σ_r can be equivalently defined using, for $r \in \mathbb{R}$, the norm $\|u\|_{\Sigma_r} := \|(1 - \Delta + |x|^2)^{r/2} u\|_{L^2}$.
- $\mathscr{S}(\mathbb{R}^3) = \bigcap_{m \ge 0} \Sigma_m$ is the space of Schwartz functions; $\mathscr{S}'(\mathbb{R}^3) = \bigcup_{m \le 0} \Sigma_m$ is the space of tempered distributions.
- We set $z_j = z_{jR} + \mathrm{i} z_{jI}$ for $z_{jR}, z_{jI} \in \mathbb{R}$.
- For $f : \mathbb{C}^n \to \mathbb{C}$, set $D_{jR} f(z) := \partial f / \partial z_{jR}(z)$ and $D_{jI} f(z) := \partial f / \partial z_{jI}(z)$.
- We set $\partial_l := \partial_{z_l}$ and $\partial_{\bar{l}} := \partial_{\bar{z}_l}$. Here, as is customary, $\partial_{z_l} = \frac{1}{2}(D_{lR} - \mathrm{i} D_{lI})$ and $\partial_{\bar{z}_l} = \frac{1}{2}(D_{lR} + \mathrm{i} D_{lI})$.
- Occasionally we use a single index $\ell = j, \bar{j}$. To define $\bar{\ell}$ we use the convention $\bar{\bar{j}} = j$. We will also write $z_{\bar{j}} = \bar{z}_j$.
- We will consider vectors $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ and, for vectors $\mu, \nu \in (\mathbb{N} \cup \{0\})^n$, we set $z^\mu \bar{z}^\nu := z_1^{\mu_1} \cdots z_n^{\mu_n} \bar{z}_1^{\nu_1} \cdots \bar{z}_n^{\nu_n}$. We will set $|\mu| = \sum_j \mu_j$.
- We have $dz_j = dz_{jR} + \mathrm{i}\, dz_{jI}$, $d\bar{z}_j = dz_{jR} - \mathrm{i}\, dz_{jI}$.
- We consider the vector $\boldsymbol{e} = (e_1, \dots, e_n)$ whose entries are the eigenvalues of H.
- P_c is the orthogonal projection of L^2 onto $\mathscr{H}_c[0]$.

- Given two Banach spaces X and Y we denote by $B(X, Y)$ the space of bounded linear operators $X \to Y$ with the norm of the uniform operator topology.

Coordinates. The first thing we need is an ansatz. This is provided by the following lemma:

Lemma 2.1. *There exist $c_0 > 0$ and $C > 0$ such that, for all $u \in H^1$ with $\|u\|_{H^1} < c_0$, there exists a unique pair $(z, \Theta) \in \mathbb{C}^n \times (H^1 \cap \mathcal{H}_c[z])$ such that*

$$u = \sum_{j=1}^{n} Q_{jz_j} + \Theta \quad \textit{with} \quad |z| + \|\Theta\|_{H^1} \le C\|u\|_{H^1}. \tag{2-1}$$

Finally, the map $u \mapsto (z, \Theta)$ is $C^\infty(B_{H^1}(0, c_0), \mathbb{C}^n \times H^1)$ and satisfies the gauge property

$$z(e^{\mathrm{i}\vartheta}u) = e^{\mathrm{i}\vartheta}z(u) \quad \textit{and} \quad \Theta(e^{\mathrm{i}\vartheta}u) = e^{\mathrm{i}\vartheta}\Theta(u). \tag{2-2}$$

Proof. We consider the functions

$$F_{jA}(u, z) := \mathrm{Re}\left\langle u - \sum_{l=1}^{n} Q_{lz_l}, \mathrm{i}\overline{D_{jA}Q_{jz_j}}\right\rangle \quad \text{for } A = R, I.$$

We have $F_{jR}(0, 0) = F_{jI}(0, 0) = 0$. These functions are smooth in $L^2 \times B_{\mathbb{C}^n}(0, b_0)$ for the b_0 in Definition 1.2. We have $F_{jR}(0, z) = \mathrm{Im}\, z_j + O(z^3)$ and $F_{jI}(0, z) = \mathrm{Re}\, z_j + O(z^3)$ by Proposition 1.1. By the implicit function theorem, there is a map $u \to z$ which is $C^\infty(B_{L^2}(0, c_0), \mathbb{C}^n)$ for $c_0 > 0$ sufficiently small. Set $\Theta := u - \sum_{j=1}^{n} Q_{jz_j}$. Then $\Theta \in C^\infty(B_{H^1}(0, c_0), H^1)$. The inequalities follow from $|z(u)| \le C\|u\|_{H^1}$, which follows from $z \in C^1$ and $z(0) = 0$. Formula (2-2) follows from

$$e^{\mathrm{i}\vartheta}u = \sum_{j=1}^{n} e^{\mathrm{i}\vartheta}Q_{jz_j} + e^{\mathrm{i}\vartheta}\Theta = \sum_{j=1}^{n} Q_{je^{\mathrm{i}\vartheta}z_j} + e^{\mathrm{i}\vartheta}\Theta$$

and from the fact that $\Theta \in \mathcal{H}_c[z]$ implies $e^{\mathrm{i}\vartheta}\Theta \in \mathcal{H}_c[z']$, where $z' = e^{\mathrm{i}\vartheta}z$. This last fact is elementary. Indeed, setting only for this proof $z_j = x_j + \mathrm{i}y_j$ and $z'_j = x'_j + \mathrm{i}y'_j$, we have

$$\mathrm{Re}\langle \mathrm{i}\overline{e^{\mathrm{i}\vartheta}\Theta}, \partial_{x'_j}Q_{jz'_j}\rangle = \partial_{x'_j}x_j\, \mathrm{Re}\langle \mathrm{i}\overline{e^{\mathrm{i}\vartheta}\Theta}, e^{\mathrm{i}\vartheta}\partial_{x_j}Q_{jz_j}\rangle + \partial_{x'_j}y_j\, \mathrm{Re}\langle \mathrm{i}\overline{e^{\mathrm{i}\vartheta}\Theta}, e^{\mathrm{i}\vartheta}\partial_{y_j}Q_{jz_j}\rangle = 0$$

if $\Theta \in \mathcal{H}_c[z]$. Similarly, $\mathrm{Re}\langle \mathrm{i}\overline{e^{\mathrm{i}\vartheta}\Theta}, \partial_{y'_j}Q_{jz'_j}\rangle = 0$. Hence $\Theta \in \mathcal{H}_c[z]$ implies $e^{\mathrm{i}\vartheta}\Theta \in \mathcal{H}_c[e^{\mathrm{i}\vartheta}z]$. □

Definition 2.2. Given $z \in \mathbb{C}^n$, we denote by $\widehat{\boldsymbol{Z}}$ the vector with entries $(z_i\bar{z}_j)$ with $i, j \in [1, n]$, in lexicographic order. We denote by $\boldsymbol{Z}$ the vector with entries $(z_i\bar{z}_j)$ with $i, j \in [1, n]$, in lexicographic order but only for pairs of indexes with $i \ne j$. Here, $\boldsymbol{Z}$ is in L, the subspace of $\mathbb{C}^{n_0} = \{(a_{i,j})_{i,j=1,\dots,n} : i \ne j\}$, $n_0 = n(n-1)$, with $(a_{i,j}) \in L$ if and only if $a_{i,j} = \bar{a}_{j,i}$ for all i, j. For a multiindex $\boldsymbol{m} = \{m_{ij} \in \mathbb{N}_0 : i \ne j\}$, we set $\boldsymbol{Z}^{\boldsymbol{m}} = \prod (z_i\bar{z}_j)^{m_{ij}}$ and $|\boldsymbol{m}| := \sum_{i,j} m_{ij}$.

We need a system of independent coordinates, which the (z, Θ) in (2-1) are not. The following lemma is used to complete the z with a continuous coordinate.

Lemma 2.3. *There exists $d_0 > 0$ such that, for any $z \in \mathbb{C}$ with $|z| < d_0$, there exists an $\mathbb{R}$-linear operator $R[z] : \mathcal{H}[0] \to \mathcal{H}_c[z]$ such that $P_c|_{\mathcal{H}_c[z]} = R[z]^{-1}$, with P_c the orthogonal projection of L^2 onto $\mathcal{H}_c[0]$; see Definition 1.2. Furthermore, for $|z| < d_0$ and $\eta \in \mathcal{H}_c[0]$, we have the following properties:*

(1) *$R[z] \in C^\infty(B_{\mathbb{C}^n}(0, d_0), B(H^1, H^1))$ with $B(H^1, H^1)$ the Banach space of $\mathbb{R}$-linear bounded operators from H^1 into itself.*

(2) *For any $r > 0$, we have $\|(R[z] - 1)\eta\|_{\Sigma_r} \le c_r |z|^2 \|\eta\|_{\Sigma_{-r}}$ for a fixed c_r.*

(3) *We have the covariance property $R[e^{i\vartheta} z] = e^{i\vartheta} R[z] e^{-i\vartheta}$.*

(4) *We have, summing on repeated indexes,*

$$R[z]\eta = \eta + (\alpha_j[z]\eta)\phi_j \quad \textit{with} \quad \alpha_j[z]\eta = \langle B_j(z), \eta\rangle + \langle C_j(z), \bar{\eta}\rangle, \tag{2-3}$$

where $B_j(z) = \widehat{B}_j(\widehat{\mathbf{Z}})$ and $C_j(z) = z_i z_\ell \widehat{C}_{i\ell j}(\widehat{\mathbf{Z}})$ for $\widehat{B}$ and $\widehat{C}_{i\ell j}$ smooth and the $\widehat{\mathbf{Z}}$ of Definition 2.2.

(5) *We have, for $r \in \mathbb{R}$ and $\mathbf{Z}$ as in Definition 2.2,*

$$\|B_j(z) + \partial_{\bar{z}_j} \bar{q}_{j z_j}\|_{\Sigma_r} + \|C_j(z) - \partial_{\bar{z}_j} q_{j z_j}\|_{\Sigma_r} \le c_r |\mathbf{Z}|^2. \tag{2-4}$$

Proof. Summing over repeated indexes, we search for a map $R[z] : L^2 \to \mathcal{H}_c[z]$ of the form

$$R[z] f = f + (\alpha_j[z] f)\phi_j \quad \text{with} \quad \alpha_j[z] f = \langle B'_j(z), f\rangle + \langle C_j(z), \bar{f}\rangle$$

such that $R[z] f \in \mathcal{H}_c[z]$ for all $f \in L^2$. The latter condition can be expressed as

$$\operatorname{Re}\langle \bar{f}, \mathrm{i} D_{lA} Q_{l z_l} + \langle \phi_j, \mathrm{i} D_{lA} Q_{l z_l}\rangle \bar{B}'_j - \langle \phi_j, \mathrm{i} D_{lA} \bar{Q}_{l z_l}\rangle C_j\rangle = 0 \quad \text{for all } f \in L^2.$$

This and the equalities

$$\begin{aligned} \langle \phi_j, \mathrm{i} D_{lR} Q_{l z_l}\rangle &= \mathrm{i}\delta_{jl} + \langle \phi_j, \mathrm{i} D_{lR} q_{l z_l}\rangle, & \langle \phi_j, \mathrm{i} D_{lI} Q_{l z_l}\rangle &= -\delta_{jl} + \langle \phi_j, \mathrm{i} D_{lI} q_{l z_l}\rangle, \\ \langle \phi_j, \mathrm{i} D_{lR} \bar{Q}_{l z_l}\rangle &= \mathrm{i}\delta_{jl} + \langle \phi_j, \mathrm{i} D_{lR} \bar{q}_{l z_l}\rangle, & \langle \phi_j, \mathrm{i} D_{lI} \bar{Q}_{l z_l}\rangle &= \delta_{jl} + \langle \phi_j, \mathrm{i} D_{lI} \bar{q}_{l z_l}\rangle, \end{aligned}$$

yield the equalities

$$\begin{aligned} D_{lR} Q_{l z_l} + (\delta_{jl} + \langle \phi_j, D_{lR} q_{l z_l}\rangle)\bar{B}'_j - (\delta_{jl} + \langle \phi_j, D_{lR} \bar{q}_{l z_l}\rangle) C_j &= 0, \\ \mathrm{i} D_{lI} Q_{l z_l} + (-\delta_{jl} + \mathrm{i}\langle \phi_j, D_{lI} q_{l z_l}\rangle)\bar{B}'_j - (\delta_{jl} + \mathrm{i}\langle \phi_j, D_{lI} \bar{q}_{l z_l}\rangle) C_j &= 0. \end{aligned}$$

They can be rewritten as

$$\begin{aligned} \phi_l + \partial_l q_{l z_l} + (\delta_{jl} + \mathrm{i}\langle \phi_j, \partial_l q_{l z_l}\rangle)\bar{B}'_j - \langle \phi_j, \partial_l \bar{q}_{l z_l}\rangle C_j &= 0, \\ \partial_{\bar{l}} q_{l z_l} + \langle \phi_j, \partial_{\bar{l}} q_{l z_l}\rangle \bar{B}'_j - (\delta_{jl} + \langle \phi_j, \partial_{\bar{l}} \bar{q}_{l z_l}\rangle) C_j &= 0. \end{aligned} \tag{2-5}$$

For $z^2 = \{z_j^2 \delta_{ij}\}$ and $\bar{z}^2 = \{\bar{z}_j^2 \delta_{ij}\}$ two $n \times n$ matrices, the solution of this system is of the form

$$\begin{pmatrix} \bar{B}' \\ C \end{pmatrix} = \sum_{m=0}^{\infty} (-1)^m \begin{pmatrix} \boldsymbol{A}_1 & \bar{z}^2 \boldsymbol{A}_2 \\ z^2 \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix}^m \begin{pmatrix} u_1 \\ z^2 u_2 \end{pmatrix}, \tag{2-6}$$

where $A_l = A_l(|z_1|^2, \dots, |z_n|^2)$ are $n \times n$ matrices and $u_l = u_l(|z_1|^2, \dots, |z_n|^2)$ are $n \times 1$ matrices for $l = 1$ (resp. $l = 2$) with entries $\phi_j + \partial_j q_{jz_j}$ (resp. $\partial_{\bar{j}} q_{jz_j}$) as $j = 1, \dots, n$. This yields the structure $\bar{B}'(z) = \widehat{B}'(\widehat{\mathbf{Z}})$ and $C_j(z) = z_i z_\ell \widehat{C}_{i\ell j}(\widehat{\mathbf{Z}})$.

Using $\langle \phi_j, q_{jz_j}\rangle = 0$, we can rewrite (2-5) in the form

$$\begin{aligned}\bar{B}'_l &= -\phi_l - \partial_l q_{lz_l} - \sum_{j\neq l} (\mathrm{i}\langle \phi_j, \partial_l q_{lz_l}\rangle \bar{B}'_j - \langle \phi_j, \partial_l \bar{q}_{lz_l}\rangle C_j),\\ C_l &= \partial_{\bar{l}} q_{lz_l} + \sum_{j\neq l} (\langle \phi_j, \partial_{\bar{l}} q_{lz_l}\rangle \bar{B}'_j - \langle \phi_j, \partial_{\bar{l}} \bar{q}_{lz_l}\rangle) C_j.\end{aligned} \tag{2-7}$$

By Proposition 1.1, this implies

$$\|\bar{B}'_l + \phi_l\|_{\Sigma_r} + \|C_l\|_{\Sigma_r} \le C|z_l|^2. \tag{2-8}$$

Reiterating this estimate, from (2-7) and for B_l defined by the following formula, we get

$$\begin{aligned}\Big\| \overbrace{\bar{B}'_l + \phi_l - \sum_{j\neq l} \mathrm{i}\langle \phi_j, \partial_l q_{lz_l}\rangle \phi_j}^{\bar{B}_l} + \partial_l q_{lz_l} \Big\|_{\Sigma_r} &\le C|\mathbf{Z}|^2\\ \|C_l - \partial_{\bar{l}} q_{lz_l}\|_{\Sigma_r} &\le C|\mathbf{Z}|^2.\end{aligned}$$

This yields (2-4). Claim (3) follows by

$$\alpha_j[e^{\mathrm{i}\vartheta} z]\eta = e^{\mathrm{i}\vartheta} \alpha_j[z] e^{-\mathrm{i}\vartheta} \eta, \tag{2-9}$$

which in turn follows by claim (4). Indeed,

$$\begin{aligned}\alpha_j[e^{\mathrm{i}\vartheta} z]\eta &= \langle \widehat{B}_j(\widehat{\mathbf{Z}}), \eta\rangle + \langle e^{2\mathrm{i}\vartheta} z_i z_\ell \widehat{C}_{i\ell j}(\widehat{\mathbf{Z}}), \bar{\eta}\rangle\\ &= e^{\mathrm{i}\vartheta}\langle \widehat{B}_j(\widehat{\mathbf{Z}}), e^{-\mathrm{i}\vartheta}\eta\rangle + e^{\mathrm{i}\vartheta}\langle z_i z_\ell \widehat{C}_{i\ell j}(\widehat{\mathbf{Z}}), \overline{e^{-\mathrm{i}\vartheta}\eta}\rangle = e^{\mathrm{i}\vartheta}\alpha_j[z]e^{-\mathrm{i}\vartheta}\eta.\end{aligned}$$

□

We are now able to define a system of coordinates near the origin in L^2.

Lemma 2.4. *For the d_0 of Lemma 2.3, the map $(z, \eta) \mapsto u$ defined by*

$$u = \sum_{j=1}^{n} Q_{jz_j} + R[z]\eta \quad \textit{for } (z, \eta) \in B_{\mathbb{C}^n}(0, d_0) \times (H^1 \cap \mathcal{H}_c[0]) \tag{2-10}$$

has values in H^1 and is C^∞. Furthermore, there is a $d_1 > 0$ such that the above map is a diffeomorphism for $(z, \eta) \in B_{\mathbb{C}^n}(0, d_1) \times (B_{H^1}(0, d_1) \cap \mathcal{H}_c[0])$ and

$$|z| + \|\eta\|_{H^1} \sim \|u\|_{H^1}. \tag{2-11}$$

Finally, we have the gauge properties $u(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} u(z, \eta)$,

$$z(e^{\mathrm{i}\vartheta} u) = e^{\mathrm{i}\vartheta} z(u) \quad \textit{and} \quad \eta(e^{\mathrm{i}\vartheta} u) = e^{\mathrm{i}\vartheta}\eta(u). \tag{2-12}$$

Proof. The smoothness follows from the smoothness in z in Proposition 1.1 and Lemma 2.3. Property $u(e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta}u(z,\eta)$ and its equivalent formula (2-12) follow from (2-2) and Lemma 2.3(3). Notice that $u = u(z,\eta)$ is the inverse of the smooth map $u \mapsto (z, \Theta) \mapsto (z, P_c\Theta)$. Formula (2-11) follows by the estimates in Proposition 1.1 and by Lemma 2.3(2). □

Resonant sets.

Definition 2.5. Consider the set of multiindexes $\boldsymbol{m}$ as in Definition 2.2 and, for any $k \in \{1, \dots, n\}$, the set

$$\begin{aligned} \mathcal{M}_k(r) &= \big\{\boldsymbol{m} : \textstyle\sum_{i=1}^n \sum_{j=1}^n m_{ij}(e_i - e_j) - e_k < 0 \text{ and } |\boldsymbol{m}| \le r\big\}, \\ \mathcal{M}_0(r) &= \big\{\boldsymbol{m} : \textstyle\sum_{i=1}^n \sum_{j=1}^n m_{ij}(e_i - e_j) = 0 \text{ and } |\boldsymbol{m}| \le r\big\}. \end{aligned} \tag{2-13}$$

Set now

$$\begin{aligned} M_k(r) &= \{(\mu,\nu) \in \mathbb{N}_0^n \times \mathbb{N}_0^n : z^\mu \bar{z}^\nu = \bar{z}_k \boldsymbol{Z}^{\boldsymbol{m}} \text{ for some } \boldsymbol{m} \in \mathcal{M}_k(r)\}, \\ M(r) &= \bigcup_{k=1}^n M_k(r) \quad \text{and} \quad M = M(2N+4) \end{aligned} \tag{2-14}$$

Lemma 2.6. *Assuming* (H3) *we have the following facts*:

(1) *If* $\boldsymbol{Z}^{\boldsymbol{m}} = z^\mu \bar{z}^\nu$, *then* $\boldsymbol{m} \in \mathcal{M}_0(2N+4)$ *implies* $\mu = \nu$. *In particular*, $\boldsymbol{m} \in \mathcal{M}_0(2N+4)$ *implies* $\boldsymbol{Z}^{\boldsymbol{m}} = |z_1|^{2l_1} \cdots |z_n|^{2l_n}$ *for some* $(l_1, \dots, l_n) \in \mathbb{N}_0^n$.

(2) *For* $|\boldsymbol{m}| \le 2N+3$ *and any* j, *we have* $\sum_{a,b}(e_a - e_b)m_{ab} - e_j \ne 0$.

Proof. First of all, if $\mu = \nu$ then $z^\mu \bar{z}^\nu = |z_1|^{2\mu_1} \cdots |z_n|^{2\mu_n}$. So the first sentence in claim (1) implies the second sentence in claim (1). We have

$$\boldsymbol{Z}^{\boldsymbol{m}} = \prod_{i,l=1}^n (z_i \bar{z}_l)^{m_{il}} = \prod_{i=1}^n z_i^{\sum_{l=1}^n m_{il}} \bar{z}_i^{\sum_{l=1}^n m_{li}} = z^\mu \bar{z}^\nu.$$

The pair (μ,ν) satisfies $|\mu| = |\nu| \le 2N+4$, by

$$|\mu| = \sum_l \mu_l = \sum_{i,l} m_{il} = |\nu|.$$

We have $(\mu - \nu) \cdot \boldsymbol{e} = 0$ by $\boldsymbol{m} \in \mathcal{M}_0(2N+4)$ and

$$\sum_i \mu_i e_i - \sum_l \nu_l e_l = \sum_{i,l} m_{il}(e_i - e_l) = 0.$$

We conclude, by (H3), that $\mu - \nu = 0$. This proves the first sentence of claim (1).

The proof of claim (2) is similar. Set

$$\boldsymbol{Z}^{\boldsymbol{m}} \bar{z}_j = \prod_{i,l=1}^n (z_i \bar{z}_l)^{m_{il}} \bar{z}_j = \prod_{i=1}^n z_i^{\sum_{l=1}^n m_{il}} \bar{z}_i^{\sum_{l=1}^n m_{li}} \bar{z}_j = z^\mu \bar{z}^\nu.$$

We have

$$(\mu - \nu) \cdot \boldsymbol{e} = \sum_i \mu_i e_i - \sum_l \nu_l e_l = \sum_{i,l} m_{il}(e_i - e_l) - e_j$$

and

$$|\mu| = \sum_{l} \mu_l = \sum_{i,l} m_{il} = |\nu| - 1. \tag{2-15}$$

If $(\mu - \nu) \cdot \boldsymbol{e} = 0$ then, by $|\mu - \nu| \leq 4N + 5$ and (H3), we would have $\mu = \nu$, which is impossible by (2-15). □

Lemma 2.7. (1) *Consider* $\boldsymbol{m} = (m_{ij}) \in \mathbb{N}_0^{n_0}$ *such that* $\sum_{i<j} m_{ij} > N$ *for* $N > |e_1|(\min\{e_j - e_i : j > i\})^{-1}$; *see* (H3). *Then, for any eigenvalue* e_k, *we have*

$$\sum_{i<j} m_{ij}(e_i - e_j) - e_k < 0. \tag{2-16}$$

(2) *Consider* $\boldsymbol{m} \in \mathbb{N}_0^{n_0}$ *with* $|\boldsymbol{m}| \geq 2N + 3$ *and the monomial* $z_j \boldsymbol{Z}^{\boldsymbol{m}}$. *Then there exist* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{N}_0^{n_0}$ *such that*

$$\sum_{i<j} a_{ij} = N + 1 = \sum_{i<j} b_{ij},$$
$$a_{ij} = b_{ij} = 0 \quad \text{for all } i > j \qquad \text{and} \qquad a_{ij} + b_{ij} \leq m_{ij} + m_{ji} \quad \text{for all } (i, j), \tag{2-17}$$

and moreover there is a pair of indexes (k, l) *such that*

$$\sum_{i<j} a_{ij}(e_i - e_j) - e_k < 0 \quad \text{and} \quad \sum_{i<j} b_{ij}(e_i - e_j) - e_l < 0 \tag{2-18}$$

and such that, for $|z| \leq 1$,

$$|z_j \boldsymbol{Z}^{\boldsymbol{m}}| \leq |z_j||z_k \boldsymbol{Z}^{\boldsymbol{a}}||z_l \boldsymbol{Z}^{\boldsymbol{b}}|. \tag{2-19}$$

(3) *For* $\boldsymbol{m}$ *with* $|\boldsymbol{m}| \geq 2N + 3$, *there exist* (k, l), $\boldsymbol{a} \in \mathcal{M}_k$ *and* $\boldsymbol{b} \in \mathcal{M}_l$ *such that* (2-19) *holds.*

Proof. Equation (2-16) follows immediately from

$$\sum_{i<j} m_{ij}(e_i - e_j) - e_k \leq -\min\{e_j - e_i : j > i\}N - e_1 < 0,$$

where the latter inequality follows by the definition of N.

Given $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{N}_0^{n_0}$ satisfying (2-17), by claim (1) they satisfy (2-18) for any pair of indexes (k, l). Consider now the monomial $z_j \boldsymbol{Z}^{\boldsymbol{m}}$. Since $|\boldsymbol{m}| \geq 2N + 3$, there are vectors $\boldsymbol{c}, \boldsymbol{d} \in \mathbb{N}_0^{n_0}$ such that $|\boldsymbol{c}| = |\boldsymbol{d}| = N + 1$ and $c_{ij} + d_{ij} \leq m_{ij}$ for all (i, j). Furthermore, we have

$$z_j \boldsymbol{Z}^{\boldsymbol{m}} = z_j z^{\mu} \bar{z}^{\nu} \boldsymbol{Z}^{\boldsymbol{c}} \boldsymbol{Z}^{\boldsymbol{d}} \quad \text{with} \quad |\mu| > 0 \text{ and } |\nu| > 0. \tag{2-20}$$

So, for z_k a factor of z^{μ} and $\bar{z}_l$ a factor of $\bar{z}^{\nu}$, and for

$$a_{ij} = \begin{cases} c_{ij} + c_{ji} & \text{for } i < j, \\ 0 & \text{for } i > j, \end{cases} \qquad b_{ij} = \begin{cases} d_{ij} + d_{ji} & \text{for } i < j, \\ 0 & \text{for } i > j, \end{cases} \tag{2-21}$$

for $|z| \leq 1$ we have, from (2-20),

$$|z_j \boldsymbol{Z}^{\boldsymbol{m}}| \leq |z_j||z_k \boldsymbol{Z}^{\boldsymbol{c}}||z_l \boldsymbol{Z}^{\boldsymbol{d}}| = |z_j||z_k \boldsymbol{Z}^{\boldsymbol{a}}||z_l \boldsymbol{Z}^{\boldsymbol{b}}|.$$

Furthermore, (2-17) is satisfied.

Since our $(\boldsymbol{a}, \boldsymbol{b})$ satisfy $\boldsymbol{a} \in \mathcal{M}_k$ and $\boldsymbol{b} \in \mathcal{M}_l$, claim (3) is a consequence of claim (2). $\square$

We end this section by exploiting the notation introduced in Lemma 2.3(5) to introduce two classes of functions. First of all, notice that the linear maps $\eta \mapsto \langle \eta, \phi_j \rangle$ extend to bounded linear maps $\Sigma_r \to \mathbb{R}$ for any $r \in \mathbb{R}$. We set

$$\Sigma_r^c := \{\eta \in \Sigma_r : \langle \eta, \phi_j \rangle = 0,\ j = 1, \dots, n\}. \tag{2-22}$$

The following two classes of functions will be used in the rest of the paper. Recall that in Definition 2.2 we introduced the space L with $\dim L = n(n-1)$. In Definitions 2.8–2.9, we denote by $\boldsymbol{Z}$ an auxiliary variable independent of z which takes values in L.

Definition 2.8. Let $\mathfrak{B}$ be an open subset of a Banach space. We will say that $F(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)$ in $C^M(I \times \mathfrak{B} \times \mathcal{A}, \mathbb{R})$, with I a neighborhood of 0 in $\mathbb{R}$ and $\mathcal{A}$ a neighborhood of 0 in $\mathbb{C}^n \times L \times \Sigma_{-K}^c$, is $F = \mathcal{R}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)$ if there exists a $C > 0$ and a smaller neighborhood $\mathcal{A}'$ of 0 such that

$$|F(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)| \le C(\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}|)^j (\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}| + |z|)^i \quad \text{in } I \times \mathfrak{B} \times \mathcal{A}'. \tag{2-23}$$

We will specify $F = \mathcal{R}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z})$ if

$$|F(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)| \le C|\boldsymbol{Z}|^j |z|^i \tag{2-24}$$

and $F = \mathcal{R}^{i,j}_{K,M}(t, \mathfrak{b}, z, \eta)$ if

$$|F(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)| \le C\|\eta\|^j_{\Sigma_{-K}} (\|\eta\|_{\Sigma_{-K}} + |z|)^i. \tag{2-25}$$

We will omit t or $\mathfrak{b}$ if there is no dependence on such variables.

We write $F = \mathcal{R}^{i,j}_{K,\infty}$ if $F = \mathcal{R}^{i,j}_{K,m}$ for all $m \ge M$. We write $F = \mathcal{R}^{i,j}_{\infty,M}$ if, for all $k \ge K$, the above F is the restriction of an $F(t, \mathfrak{b}, z, \eta) \in C^M(I \times \mathfrak{B} \times \mathcal{A}_k, \mathbb{R})$ with $\mathcal{A}_k$ a neighborhood of 0 in $\mathbb{C}^n \times L \times \Sigma_{-k}^c$ and which is $F = \mathcal{R}^{i,j}_{k,M}$. Finally we write $F = \mathcal{R}^{i,j}_{\infty,\infty}$ if $F = \mathcal{R}^{i,j}_{k,\infty}$ for all k.

Definition 2.9. We will say that $T(t, \mathfrak{b}, z, \eta) \in C^M(I \times \mathfrak{B} \times \mathcal{A}, \Sigma_K(\mathbb{R}^3, \mathbb{C}))$, with the above notation, is $T = \boldsymbol{S}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)$ if there exists a $C > 0$ and a smaller neighborhood $\mathcal{A}'$ of 0 such that

$$\|T(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)\|_{\Sigma_K} \le C(\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}|)^j (\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}| + |z|)^i \quad \text{in } I \times \mathfrak{B} \times \mathcal{A}'. \tag{2-26}$$

We use notations $\boldsymbol{S}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z})$, $\boldsymbol{S}^{i,j}_{K,M}(t, \mathfrak{b}, z, \eta)$, etc. as above.

Notice that we have the elementary formulas

$$\mathcal{R}^{a,b}_{K,M} \boldsymbol{S}^{i,j}_{K,M} = \boldsymbol{S}^{i+a,j+b}_{K,M} \quad \text{and} \quad \mathcal{R}^{a,b}_{K,M} \mathcal{R}^{i,j}_{K,M} = \mathcal{R}^{i+a,j+b}_{K,M}. \tag{2-27}$$

Remark 2.10. For functions $F(t, \mathfrak{b}, z, \eta)$ and $T(t, \mathfrak{b}, z, \eta)$, we write $F(t, \mathfrak{b}, z, \eta) = \mathcal{R}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)$ and $T(t, \mathfrak{b}, z, \eta) = \boldsymbol{S}^{i,j}_{K,M}(t, \mathfrak{b}, z, \boldsymbol{Z}, \eta)$ when the equality holds restricting the variable $\boldsymbol{Z}$ to the $\boldsymbol{Z}$ of Definition 2.2 for symbols satisfying Definitions 2.8–2.9.

Furthermore, later, when we write $\mathcal{R}^{i,j}_{K,M}$ and $\boldsymbol{S}^{i,j}_{K,M}$, we will mean $\mathcal{R}^{i,j}_{K,M}(z, \boldsymbol{Z}, \eta)$ and $\boldsymbol{S}^{i,j}_{K,M}(z, \boldsymbol{Z}, \eta)$, respectively.

Notice that $F = \mathcal{R}^{i,j}_{K,M}(z, \boldsymbol{Z})$ or $T = \boldsymbol{S}^{i,j}_{K,M}(z, \boldsymbol{Z})$ do not mean independence from the variable η.

3. Invariants

Equation (1-1) admits energy and mass invariants, defined as follows:

$$E(u) := E_K(u) + E_P(u), \quad \text{where} \quad E_K(u) := \langle Hu, \bar{u}\rangle \text{ and } E_P(u) = \frac{1}{2}\int_{\mathbb{R}^3} |u(x)|^4\,dx, \tag{3-1}$$
$$Q(u) := \langle u, \bar{u}\rangle.$$

We have $E \in C^\infty(H^1(\mathbb{R}^3, \mathbb{C}), \mathbb{R})$ and $Q \in C^\infty(L^2(\mathbb{R}^3, \mathbb{C}), \mathbb{R})$. We denote by dE the Fréchet derivative of E. We define $\nabla E \in C^\infty(H^1(\mathbb{R}^3, \mathbb{C}), H^{-1}(\mathbb{R}^3, \mathbb{C}))$ by $dEX = \mathrm{Re}\langle \nabla E, \overline{X}\rangle$ for any $X \in H^1$. We define also $\nabla_u E$ and $\nabla_{\bar{u}} E$ by

$$dEX = \langle \nabla_u E, X\rangle + \langle \nabla_{\bar{u}} E, \overline{X}\rangle, \quad \text{that is,} \quad \nabla_u E = 2^{-1}\overline{\nabla E} \text{ and } \nabla_{\bar{u}} E = 2^{-1}\nabla E.$$

Notice that $\nabla E = 2Hu + 2|u|^2 u$. Then (1-1) can be interpreted as

$$\mathrm{i}\dot{u} = \nabla_{\bar{u}} E(u). \tag{3-2}$$

Lemma 3.1. *Consider the coordinates $(z, \eta) \mapsto u$ in Lemma 2.4. Then there exists some functions as in Definitions 2.8–2.9 such that, for $(z, \eta) \in B_{\mathbb{C}^n}(0, d_0) \times (B_{H^1}(0, d_0) \cap \mathscr{H}_c[0])$, we have, for any preassigned $r_0 \in \mathbb{N}$, the expansion (where* c.c. *means complex conjugate)*

$$\begin{aligned}
E(u) = {} & \sum_{j=1}^n E(Q_{jz_j}) + \langle H\eta, \bar{\eta}\rangle + \mathscr{R}^{1,2}_{r_0,\infty}(z, \eta) \\
& + \sum_{j\neq k}\big[E_{jz_j}(\mathrm{Re}\langle q_{jz_j}, \bar{z}_k\phi_k\rangle + \mathrm{Re}\langle q_{kz_k}, \bar{z}_j\phi_j\rangle) + \mathrm{Re}\langle |Q_{kz_k}|^2 Q_{kz_k}, \bar{z}_j\phi_j\rangle\big] \\
& + \mathscr{R}^{0,2N+5}_{r_0,\infty}(z, \boldsymbol{Z}) + \sum_{j=1}^n \sum_{l=1}^{2N+3} \sum_{|\boldsymbol{m}|=l+1} \boldsymbol{Z}^{\boldsymbol{m}} a_{j\boldsymbol{m}}(|z_j|^2) + \mathrm{Re}\langle \boldsymbol{S}^{0,2N+4}_{r_0,\infty}(z, \boldsymbol{Z}), \bar{\eta}\rangle \\
& + \sum_{j,k=1}^n \sum_{l=1}^{2N+3} \sum_{|\boldsymbol{m}|=l} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{jk\boldsymbol{m}}(|z_k|^2), \eta\rangle + \text{c.c.}) + \sum_{i+j=2} \sum_{|\boldsymbol{m}|\le 1} \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{2\boldsymbol{m}ij}(z), \eta^i \bar{\eta}^j\rangle \\
& + \sum_{d+c=3} \sum_{i+j=d} \langle G_{dij}(z), \eta^i \bar{\eta}^j\rangle \mathscr{R}^{0,c}_{r_0,\infty}(z, \eta) + E_P(\eta),
\end{aligned} \tag{3-3}$$

where:

- $(a_{j\boldsymbol{m}}, G_{jk\boldsymbol{m}}) \in C^\infty(B_{\mathbb{R}}(0, d_0), \mathbb{C} \times \Sigma_{r_0}(\mathbb{R}^3, \mathbb{C}))$;
- $(G_{2\boldsymbol{m}ij}, G_{dij}) \in C^\infty(B_{\mathbb{C}^n}(0, d_0), \Sigma_{r_0}(\mathbb{R}^3, \mathbb{C}) \times \Sigma_{r_0}(\mathbb{R}^3, \mathbb{C}))$;
- *for $|\boldsymbol{m}| = 0$, where, in particular, $G_{2\boldsymbol{0}ij}(0) = 0$, we have*
$$\sum_{i+j=2} \langle G_{2\boldsymbol{0}ij}(z), \eta^i \bar{\eta}^j\rangle = \sum_{j=1}^n \langle |Q_{jz_j}|^2 \eta, \bar{\eta}\rangle + 2\sum_{j=1}^n \mathrm{Re}\langle Q_{jz_j}\,\mathrm{Re}(Q_{jz_j}\bar{\eta}), \bar{\eta}\rangle; \tag{3-4}$$
- $\mathscr{R}^{1,2}_{r_0,\infty}(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = \mathscr{R}^{1,2}_{r_0,\infty}(z, \eta)$ *for all $\vartheta \in \mathbb{R}$ for the third term in the right-hand side of (3-3).*

Remark 3.2. In (3-3) the terms of the second line could potentially derail our proof. They appear in (3-7)–(3-9). Similarly problematic is the first term in the right-hand side in (4-18) later. All these terms are tied up. Indeed, in Lemma 4.11 we will show that in a system of coordinates better suited to search for an effective Hamiltonian the problematic terms in the expansion of E cancel out.

In the proof of Lemma 3.1 we use the following lemma:

Lemma 3.3. *We have, for $j \neq k$ and $\delta E_{jz_j} := E_{jz_j} - e_j$,*

$$E_{jz_j}\langle q_{kz_k}, \phi_j\rangle + \langle |Q_{kz_k}|^2 Q_{kz_k}, \phi_j\rangle = E_{kz_k}\langle q_{kz_k}, \phi_j\rangle + \delta E_{jz_j}\langle q_{kz_k}, \phi_j\rangle. \tag{3-5}$$

Proof. We apply $\langle \cdot , \phi_j\rangle$ to

$$Hq_{kz_k} + |Q_{kz_k}|^2 Q_{kz_k} = z_k \delta E_{kz_k}\phi_k + E_{kz_k} q_{kz_k}$$

to get the following equality, which, from $e_j = E_{jz_j} - \delta E_{jz_j}$, yields (3-5):

$$e_j\langle q_{kz_k}, \phi_j\rangle + \langle |Q_{kz_k}|^2 Q_{kz_k}, \phi_j\rangle = E_{kz_k}\langle q_{kz_k}, \phi_j\rangle. \qquad \square$$

Proof of Lemma 3.1. First of all, we have the Taylor expansion

$$E(u) = E\Big(\sum_{j=1}^n Q_{jz_j}\Big) + \operatorname{Re}\Big\langle \nabla E\Big(\sum_{j=1}^n Q_{jz_j}\Big), \overline{R[z]\eta}\Big\rangle + 2^{-1}\operatorname{Re}\Big\langle \nabla^2 E\Big(\sum_{j=1}^n Q_{jz_j}\Big) R[z]\eta, \overline{R[z]\eta}\Big\rangle + E_3(\eta) \tag{3-6}$$

with

$$E_3(\eta) := \int_0^1 (1-t)\operatorname{Re}\Big\langle\Big[\nabla^2 E_P\Big(\sum_{j=1}^n Q_{jz_j} + tR[z]\eta\Big) - \nabla^2 E_P\Big(\sum_{j=1}^n Q_{jz_j}\Big)\Big] R[z]\eta, \overline{R[z]\eta}\Big\rangle dt.$$

Step 1. We consider the expansion of the first term in the right-hand side of (3-6). We have

$$\begin{aligned}\Big|\sum Q_{jz_j}\Big|^4 = \sum |Q_{jz_j}|^4 + 4\sum_{j\neq k} |Q_{jz_j}|^2 \operatorname{Re}(Q_{jz_j}\bar Q_{kz_k}) \\ + 2\sum_{j<k} |Q_{jz_j}|^2 |Q_{kz_k}|^2 + \sum_{\substack{j\neq k\\ j'\neq k'}} \operatorname{Re}(Q_{jz_j}\bar Q_{kz_k})\operatorname{Re}(Q_{j'z_{j'}}\bar Q_{k'z_{k'}}) + 4\sum_{\substack{k<l\\ j\neq k,l}} |Q_{jz_j}|^2 \operatorname{Re}(Q_{kz_k}\bar Q_{lz_l}).\end{aligned}$$

All terms are invariant under the change of variable $z \rightsquigarrow e^{\mathrm{i}\vartheta} z$. The second line is $O(|\mathbf{Z}|^2)$. We conclude that

$$\begin{aligned} E\Big(\sum_{j=1,\dots,n} Q_{jz_j}\Big) &= \sum_{j,k}\langle HQ_{jz_j}, \bar Q_{kz_k}\rangle + \frac12 \int\Big|\sum_{j=1,\dots,n} Q_{jz_j}\Big|^4 \\ &= \sum_{j=1,\dots,n} E(Q_{jz_j}) + R_1 + \sum_{j\neq k}\big[\operatorname{Re}\langle HQ_{jz_j}, \bar Q_{kz_k}\rangle + 2\operatorname{Re}\langle |Q_{jz_j}|^2 Q_{jz_j}, \bar Q_{kz_k}\rangle\big], \end{aligned} \tag{3-7}$$

where

$$R_1 := \sum_{j<k} \int |Q_{jz_j}|^2 |Q_{kz_k}|^2 + \frac{1}{2} \sum_{\substack{j\neq k \\ j'\neq k'}} \int \mathrm{Re}(Q_{jz_j} \bar{Q}_{kz_k}) \,\mathrm{Re}(Q_{j'z_{j'}} \bar{Q}_{k'z_{k'}}) + 2 \sum_{\substack{k<l \\ j\neq k,l}} \int |Q_{jz_j}|^2 \,\mathrm{Re}(Q_{kz_k} \bar{Q}_{lz_l})$$
$$= O(|\mathbf{Z}|^2).$$

By Proposition 1.1 and by (3-5), the second summation in the last line of (3-7) equals

$$\sum_{j\neq k} [E_{jz_j} \,\mathrm{Re}\langle Q_{jz_j}, \bar{Q}_{kz_k}\rangle + \mathrm{Re}\langle |Q_{jz_j}|^2 Q_{jz_j}, \bar{Q}_{kz_k}\rangle]$$
$$= \sum_{j\neq k} \big[E_{jz_j} (\mathrm{Re}\langle q_{jz_j}, \bar{z}_k \phi_k\rangle + \mathrm{Re}\langle q_{kz_k}, \bar{z}_j \phi_j\rangle) + \mathrm{Re}\langle |Q_{kz_k}|^2 Q_{kz_k}, \bar{z}_j \phi_j\rangle\big] + R_2, \quad (3\text{-}8)$$

where

$$R_2 := \sum_{j\neq k} E_{jz_j} \,\mathrm{Re}\langle q_{jz_j}, \bar{q}_{kz_k}\rangle + \mathrm{Re}\langle |Q_{kz_k}|^2 Q_{kz_k}, \bar{q}_{jz_j}\rangle = O(|\mathbf{Z}|^2).$$

The summation in (3-8) is $O(|z|^2 |\mathbf{Z}|)$ and not of the form $O(|\mathbf{Z}|^2)$. Indeed, in the particular case when $z_k = \rho_k$ and $z_j = \rho_j$ are real numbers, we have what follows, which is not $O(\rho_k^2 \rho_j^2)$:

$$E_{jz_j} \,\mathrm{Re}\langle q_{jz_j}, \bar{z}_k \phi_k\rangle + E_{kz_k} \,\mathrm{Re}\langle q_{kz_k}, \bar{z}_j \phi_j\rangle + \mathrm{Re}\langle |Q_{kz_k}|^2 Q_{kz_k}, \bar{z}_j \phi_j\rangle$$
$$= \rho_k \rho_j \big[E_{j\rho_j} \rho_j^2 \langle \tilde{q}_j(\rho_j^2), \phi_k\rangle + E_{k\rho_k} \rho_k^2 \langle \tilde{q}_k(\rho_k), \phi_j\rangle + \rho_k^2 \langle (\phi_k + \hat{q}_k(\rho_k^2))^3, \phi_j\rangle\big]. \quad (3\text{-}9)$$

Finally, we observe that $R_1 + R_2 = O(|\mathbf{Z}|^2)$ summed up together yield the first two terms on the third line of (3-3).

Indeed, since $R_1 + R_2$ is gauge invariant, by Lemma B.3 in Appendix B we have

$$R_1 + R_2 = \sum_{j=1}^{n} \sum_{l=1}^{2N+3} \sum_{|\boldsymbol{m}|=l+1} \mathbf{Z}^{\boldsymbol{m}} b_{j\boldsymbol{m}}(|z_j|^2) + O(|\mathbf{Z}|^{2N+5}) \quad (3\text{-}10)$$

with $O(|\mathbf{Z}|^{2N+5})$ smooth in z, independent of η and gauge invariant.

We have discussed the contribution to (3-3) of the first term in the expansion (3-6). Now we consider the other terms in (3-6).

Step 2. We consider the expansion of the second term in the right-hand side of (3-6).

By $\mathrm{Re}\langle \nabla E(Q_{jz_j}), \overline{R[z]\eta}\rangle = 2\,\mathrm{Re}\, E_{jz_j} \langle Q_{jz_j}, \overline{R[z]\eta}\rangle = 0$, which follows from $R[z]\eta \in \mathcal{H}_c[z]$ and $\mathrm{i}Q_{jz_j} = -z_{jI} D_{jR} Q_{jz_j} + z_{jR} D_{jI} Q_{jz_j}$ — see (11) in [Gustafson et al. 2004] (and which is an immediate consequence of $Q_{jz_j} = e^{\mathrm{i}\theta} Q_{j|z_j|}$ for $z_j = e^{\mathrm{i}\theta}|z_j|$) — we have

$$\mathrm{Re}\Big\langle \nabla E\Big(\sum_{j=1}^{n} Q_{jz_j}\Big), \overline{R[z]\eta}\Big\rangle$$
$$= \overbrace{\mathrm{Re}\langle \nabla E(Q_{1z_1}), \overline{R[z]\eta}\rangle}^{0} + \int_0^1 \partial_t \,\mathrm{Re}\Big\langle \nabla E\Big(Q_{1z_1} + t \sum_{j>1} Q_{jz_j}\Big), \overline{R[z]\eta}\Big\rangle dt$$

$$= \mathrm{Re}\Big\langle \nabla E\Big(\sum_{j>1} Q_{jz_j}\Big), \overline{R[z]\eta}\Big\rangle + \int_{[0,1]^2} \partial_s\partial_t \,\mathrm{Re}\Big\langle \nabla E_P\Big(sQ_{1z_1} + t\sum_{l>1} Q_{lz_l}\Big), \overline{R[z]\eta}\Big\rangle\,dt\,ds$$

$$= \sum_{j=1}^{n-1}\int_{[0,1]^2} \partial_s\partial_t \,\mathrm{Re}\Big\langle \nabla E_P\Big(sQ_{jz_j} + t\sum_{l>j} Q_{lz_l}\Big), \overline{R[z]\eta}\Big\rangle\,dt\,ds, \tag{3-11}$$

where the last line is obtained by repeating the argument in the first three lines. For $\widehat{Q}_j = \sum_{l>j} Q_{lz_l}$, by $\nabla E_P(u) = 2|u|^2u$, the last line of (3-11) is, in the notation of Lemma 2.3,

$$2\sum_{j=1}^{n-1}\mathrm{Re}\big\langle 2Q_{jz_j}|\widehat{Q}_j|^2 + 2|Q_{jz_j}|^2\widehat{Q}_j + Q_{jz_j}^2\overline{\widehat{Q}}_j + \overline{Q}_{jz_j}\widehat{Q}_j^2,\ \bar\eta + \phi_j\big(\langle\overline{\widehat{B}}_j(\widehat{\boldsymbol{Z}}),\bar\eta\rangle + \langle\bar z_i\bar z_\ell\overline{\widehat{C}}_{i\ell j}(\widehat{\boldsymbol{Z}}),\eta\rangle\big)\big\rangle.$$

Further expanding $\widehat{Q}_j = \sum_{l>j} Q_{lz_l}$ and using $Q_{lz_l} = z_l(\phi_l + \hat q_l(|z_l|^2))$, the above term is of the form

$$\sum_{j=1}^{n}\sum_{|\boldsymbol{m}|=1}\big(\bar z_j\boldsymbol{Z}^{\boldsymbol{m}}\langle G_{j\boldsymbol{m}}(\widehat{\boldsymbol{Z}}),\eta\rangle + \text{c.c.}\big).$$

As in Step 1, by Lemma B.4, this can be expanded into

$$\sum_{j=1}^{n}\sum_{1\le|\boldsymbol{m}|\le 2N+3}\big(\bar z_j\boldsymbol{Z}^{\boldsymbol{m}}\langle G_{jk\boldsymbol{m}}(|z_k|^2),\eta\rangle + \text{c.c.}\big) + \sum_{|\boldsymbol{m}|=2N+4}\big(\boldsymbol{Z}^{\boldsymbol{m}}\langle G_{\boldsymbol{m}}(z),\eta\rangle + \text{c.c.}\big). \tag{3-12}$$

Thus the last line in (3-11) can be absorbed in the third and fourth lines of (3-3).

Step 3. We consider the expansion of the third term in the right-hand side of (3-6). Using $\nabla^2E_K(u) = 2H$ and proceeding as for (3-6), we obtain

$$2^{-1}\,\mathrm{Re}\Big\langle \nabla^2E\Big(\sum_{j=1}^{n} Q_{jz_j}\Big)R[z]\eta, \overline{R[z]\eta}\Big\rangle$$

$$= 2^{-1}\,\mathrm{Re}\Big\langle \nabla^2E_K\Big(\sum_{j=1}^{n} Q_{jz_j}\Big)R[z]\eta, \overline{R[z]\eta}\Big\rangle + 2^{-1}\sum_{j=1}^{n}\mathrm{Re}\langle\nabla^2E_P(Q_{jz_j})R[z]\eta, \overline{R[z]\eta}\rangle$$

$$+ 2^{-1}\sum_{j=1}^{n-1}\int_{[0,1]^2}\partial_s\partial_t\,\mathrm{Re}\Big\langle\nabla^2E_P\Big(sQ_{jz_j} + t\sum_{l=j+1}^{n} Q_{lz_l}\Big)R[z]\eta, \overline{R[z]\eta}\Big\rangle\,dt\,ds.$$

The third line is absorbed in the $\boldsymbol{Z}^{\boldsymbol{m}}\langle G_{2\boldsymbol{m}ij}(z),\eta^i\bar\eta^j\rangle + \mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)$ with $|\boldsymbol{m}| = 1$ terms in (3-3). From the second line, using (2-3)–(2-4) and, in particular, $\alpha_j[z]\eta = \mathcal{R}^{1,1}_{r_0,\infty}(z,\eta)$ for the last equality, we have

$$2^{-1}\,\mathrm{Re}\Big\langle \nabla^2E_K\Big(\sum_{j=1}^{n} Q_{jz_j}\Big)R[z]\eta, \overline{R[z]\eta}\Big\rangle = \langle HR[z]\eta, \overline{R[z]\eta}\rangle$$

$$= \langle H\eta,\bar\eta\rangle + 2\sum_{j=1}^{n}\mathrm{Re}\big[(\alpha_j[z]\eta)\langle H\phi_j,\bar\eta\rangle\big] + \sum_{j,k=1}^{n} e_j|\alpha_j[z]\eta|^2$$

$$= \langle H\eta,\bar\eta\rangle + \mathcal{R}^{1,2}_{r_0,\infty}(z,\eta),$$

which yield the second and third terms in the right-hand side of (3-3). For

$$2^{-1}\sum_{j=1}^{n}\nabla^2 E_P(Q_{jz_j})\eta=\sum_{j=1}^{n}|Q_{jz_j}|^2\eta+2\sum_{j=1}^{n}Q_{jz_j}\operatorname{Re}(Q_{jz_j}\overline{\eta}),$$

we have, for $G_{2\mathbf{0}ij}(z)$ as in (3-4),

$$2^{-1}\sum_{j=1}^{n}\operatorname{Re}\langle\nabla^2 E_P(Q_{jz_j})R[z]\eta,\overline{R[z]\eta}\rangle=\mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)+\sum_{i+j=2}\langle G_{2\mathbf{0}ij}(z),\eta^i\,\overline{\eta}^j\rangle. \tag{3-13}$$

This $\mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)$ defines the third term in the right-hand side of (3-3). Notice that $\mathcal{R}^{1,2}_{r_0,\infty}(e^{\mathrm{i}\vartheta}z,e^{\mathrm{i}\vartheta}\eta)=\mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)$ because this invariance is satisfied both by the left-hand side of (3-13) (by the invariance of E, (2-2) and Lemma 2.3) and by the last summation in the right-hand side of (3-13), by formula (3-4).

Step 4. We now turn to the $E_3(\eta)$ term in (3-6). By elementary computations,

$$\begin{aligned}E_3(\eta)&=\int_{[0,1]^2}t(1-t)\,d^3E_P\Big(\sum_{j\geq 1}Q_{jz_j}+stR[z]\eta\Big)\cdot(R[z]\eta)^3\,dt\,ds\\&=E_P(R[z]\eta)\\&\qquad+\int_{[0,1]^3}t(1-t)\,d^4E_P\Big(\tau\sum_{j\geq 1}Q_{jz_j}+stR[z]\eta\Big)\cdot(R[z]\eta)^3\sum_{j\geq 1}Q_{jz_j}\,dt\,ds\,d\tau\end{aligned} \tag{3-14}$$

with $d^3E_P(u)\cdot v^3$ the trilinear differential form applied to (v,v,v) and $d^4E_P(u)\cdot v^3w$ the 4-linear differential form applied to (v,v,v,w).

In particular, we have used the fact that, since $d^jE_P(0)=0$ for $0\leq j\leq 2$, we have

$$E_P(R[z]\eta)=\int_{[0,1]^2}t(1-t)\,d^3E_P(stR[z]\eta)\cdot(R[z]\eta)^3\,dt\,ds. \tag{3-15}$$

For $\beta(u)=|u|^4$, and using the fact that $d^4\beta(u)\in B^4(\mathbb{C},R)$ is constant in u, the last line of (3-14) is

$$\frac{1}{12}\int_{\mathbb{R}^3}d^4\beta\cdot((R[z]\eta)(x))^3\sum_{j\geq 1}Q_{jz_j}(x)\,dx,$$

and can be absorbed in the $\langle G_{dij}(z),\eta^i\,\overline{\eta}^j\rangle\mathcal{R}^{0,c}_{r_0,\infty}(z,\eta)$ terms in (3-3). We expand $E_P(R[z]\eta)$ as a sum of similar terms and of $E_P(\eta)$. □

In order to extract from the functional in (3-3) an effective Hamiltonian well suited for the FGR and dispersive estimates, we need to implement a Birkhoff normal form argument; see Section 5. This requires an intermediate change of coordinates, which will partially normalize the symplectic form Ω defined in (4-1) below, and diagonalize the homological equations. Notice that, as a bonus, this change of coordinates erases the bad terms in the expansion of E in (3-3) discussed in Remark 3.2.

4. Darboux theorem

System (3-2) is Hamiltonian with respect to the symplectic form in $H^1(\mathbb{R}^3, \mathbb{C})$,

$$\Omega(X, Y) := \mathrm{i}\langle X, \overline{Y}\rangle - \mathrm{i}\langle \overline{X}, Y\rangle = 2\,\mathrm{Im}\langle \overline{X}, Y\rangle. \tag{4-1}$$

In terms of the spectral decomposition of H (recall $\overline{\phi}_j = \phi_j$),

$$X = \sum_{j=1}^{n} \langle X, \phi_j\rangle \phi_j + P_c X, \tag{4-2}$$

$$\Omega(X, Y) = \mathrm{i}\sum_{j=1}^{n} (\langle X, \phi_j\rangle\langle \overline{Y}, \phi_j\rangle - \langle \overline{X}, \phi_j\rangle\langle Y, \phi_j\rangle) + \mathrm{i}\langle P_c X, P_c \overline{Y}\rangle - \mathrm{i}\langle P_c \overline{X}, P_c Y\rangle. \tag{4-3}$$

However, in terms of the coordinates in Lemma 2.4, Ω admits a quite more complicated representation, as we shall see. This will require us to adjust these coordinates.

Our first observation is that, for the coordinates in Lemma 2.4, we have the following facts:

Lemma 4.1. *The Fréchet derivatives of $\eta(u)$ and z_j are given by the formulas*

$$d\eta(u) = -\sum_{j=1,\dots,n}\sum_{A=I,R} P_c D_{jA} q_{jz_j}\, dz_{jA} + P_c, \tag{4-4}$$

$$dz_j = \langle\,\cdot\,, \phi_j\rangle - \sum_{k:k\neq j}\sum_{A=I,R} \langle D_{kA} q_{kz_k}, \phi_j\rangle\, dz_{kA} - \sum_{k=1}^{n}\sum_{A=I,R} D_{kA}\alpha_j[z]\eta\, dz_{kA} - \alpha_j[z]\circ d\eta. \tag{4-5}$$

Analogous formulas for dz_{jR} and dz_{jI} are obtained by applying Re *and* Im *to (4-5).*

Proof. We start with (4-4). By the independence of z and η, we have

$$d\eta\,\frac{\partial}{\partial z_{jR}} = d\eta\,\frac{\partial}{\partial z_{jI}} = 0, \tag{4-6}$$

where

$$\frac{\partial}{\partial z_{jA}} = D_{jA} Q_{jz_j} + \sum_{k=1}^{n} D_{jA}(\alpha_k[z]\eta)\phi_k. \tag{4-7}$$

Next, for $\xi \in \mathcal{H}_c[0]$ we have what follows, which implies $d\eta\, R[z]P_c = 1|_{\mathcal{H}_c[0]}$:

$$d\eta\, R[z]P_c\xi = \frac{d}{dt}\eta(Q_{jz_j} + R[z](\eta + t\xi))\Big|_{t=0} = \xi.$$

So $d\eta = \sum(a_j\, dz_{jR} + b_j\, dz_{jI}) + P_c$, where we used $P_c R[z] = 1$. Then a_j and b_j can be computed applying $\sum(a_j\, dz_{jR} + b_j\, dz_{jI}) + P_c$ to the vectors (4-7) and using (4-6). Finally (4-5) follows by

$$z_j(u) = \Big\langle u - \sum_{k=1}^{n} q_{kz_k} - R[z]\eta, \phi_j\Big\rangle = \Big\langle u - \sum_{k:k\neq j} q_{kz_k}, \phi_j\Big\rangle - \alpha_j[z]\eta. \qquad \square$$

We consider the function $\overline{\eta}(u)$. Notice that $d\overline{\eta}(u)X = d\overline{\eta}(u+tX)/dt\big|_{t=0} = \overline{d\eta(u)X}$. Now we introduce a new symplectic form. Notice that our final choice of symplectic form is not the Ω_0' defined here in (4-8), but rather the Ω_0 defined in (4-13).

Lemma 4.2. *Set*

$$\begin{aligned} \Omega_0' &:= 2\sum_{j=1}^{n} dz_{jR}\wedge dz_{jI} + \mathrm{i}\langle d\eta, d\overline{\eta}\rangle - \mathrm{i}\langle d\overline{\eta}, d\eta\rangle \\ \textit{and}\quad B_0' &:= \sum_{j=1}^{n}(z_{jR}\,dz_{jI} - z_{jI}\,dz_{jR}) - \frac{\mathrm{i}}{2}(\langle\overline{\eta}, d\eta\rangle - \langle\eta, d\overline{\eta}\rangle). \end{aligned} \tag{4-8}$$

Then $dB_0' = \Omega_0'$ *and* $\Omega = \Omega_0'$ *at* $u=0$ *for the* Ω *of* (4-1). *Furthermore,*

$$\Phi^* B_0' = B_0' \quad \textit{for}\ \ \Phi(u) = e^{\mathrm{i}\vartheta}u\ \ \textit{for any fixed}\ \vartheta\in\mathbb{R}. \tag{4-9}$$

Proof. The equality $dB_0' = \Omega_0'$ is elementary. Indeed, $d(z_{jR}\,dz_{jI} - z_{jI}\,dz_{jR}) = 2\,dz_{jR}\wedge dz_{jI}$ and, for a pair of constant vector fields X and Y, since $d^2\eta(X,Y) = d^2\eta(Y,X)$ we have

$$d\langle\overline{\eta}, d\eta\rangle(X,Y) = X\langle\overline{\eta}, d\eta\,Y\rangle - Y\langle\overline{\eta}, d\eta\,X\rangle = \langle d\overline{\eta}\,X, d\eta\,Y\rangle - \langle d\overline{\eta}\,Y, d\eta\,X\rangle.$$

This yields $d\langle\overline{\eta}, d\eta\rangle = \langle d\overline{\eta}, d\eta\rangle - \langle d\eta, d\overline{\eta}\rangle$ and also $d\langle\eta, d\overline{\eta}\rangle = -d\langle\overline{\eta}, d\eta\rangle = \langle d\eta, d\overline{\eta}\rangle - \langle d\overline{\eta}, d\eta\rangle$.

To compute Ω_0' at $u=0$, we observe that, by Lemma 4.1, we have $d\eta = P_c$ at $u=0$, so that

$$\mathrm{i}\langle d\eta\,X, d\overline{\eta}\,Y\rangle - \mathrm{i}\langle d\overline{\eta}\,X, d\eta\,Y\rangle = \mathrm{i}\langle P_cX, P_c\overline{Y}\rangle - \mathrm{i}\langle P_c\overline{X}, P_cY\rangle \quad \text{at}\ \ u=0. \tag{4-10}$$

By Lemma 4.1 and Proposition 1.1, at $u=0$ we have $dz_{jR} = \mathrm{Re}\langle\,\cdot\,,\phi_j\rangle$ and $dz_{jI} = \mathrm{Im}\langle\,\cdot\,,\phi_j\rangle$. Summing on repeated indexes, we have

$$\begin{aligned} \mathrm{i}(\langle X,\phi_j\rangle\langle\overline{Y},\phi_j\rangle - \langle\overline{X},\phi_j\rangle\langle Y,\phi_j\rangle) &= -2\,\mathrm{Im}(\langle X,\phi_j\rangle\langle\overline{Y},\phi_j\rangle) \\ &= 2(\mathrm{Re}\langle X,\phi_j\rangle\,\mathrm{Im}\langle Y,\phi_j\rangle - \mathrm{Re}\langle Y,\phi_j\rangle\,\mathrm{Im}\langle X,\phi_j\rangle) \\ &= 2\,\mathrm{Re}\langle\,\cdot\,,\phi_j\rangle\wedge\mathrm{Im}\langle\,\cdot\,,\phi_j\rangle(X,Y) \\ &= 2\,dz_{jR}\wedge dz_{jI}|_{u=0}(X,Y). \end{aligned} \tag{4-11}$$

By (4-10)–(4-11), we get $\Omega = \Omega_0'$ at $u=0$. Finally, (4-9) follows immediately by

$$B_0' := \sum_{j=1}^{n}\mathrm{Im}(\bar{z}_j dz_j) + \mathrm{Im}\langle\overline{\eta}, d\eta\rangle. \tag{4-12}$$

This concludes the proof. ◻

Summing on repeated indexes and using the notation in Proposition 1.1, we introduce the differential forms

$$\begin{aligned} &\Omega_0 := \Omega_0' + \mathrm{i}\gamma_j(|z_j|^2)\,dz_j\wedge d\bar{z}_j, \\ \text{where}\quad &\gamma_j(|z_j|^2) := \langle\hat{q}_j(|z_j|^2), \hat{q}_j(|z_j|^2)\rangle + 2|z_j|^2\langle\hat{q}_j(|z_j|^2), \hat{q}_j'(|z_j|^2)\rangle, \\ \text{and}\quad &B_0 := B_0' - \mathrm{Im}\langle D_{jA}\bar{q}_{jz_j}, q_{jz_j}\rangle\,dz_{jA} \end{aligned} \tag{4-13}$$

with $\hat{q}_j'(t) = d\hat{q}_j/dt$. We have the following lemma:

Lemma 4.3. *We have* $\gamma_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}(|z_j|^2)$. *We have* $dB_0 = \Omega_0$ *and*

$$\Phi^* B_0 = B_0 \quad \text{for } \Phi(u) = e^{\mathrm{i}\vartheta}u \text{ for any fixed } \vartheta \in \mathbb{R}. \tag{4-14}$$

Proof. The identity $\gamma_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}(|z_j|^2)$ is elementary from Proposition 1.1 and Definition 2.8. Next, $dB_0 = \Omega_0$ follows by $dB_0' = \Omega_0'$ and

$$\begin{aligned} -d\,\mathrm{Im}\langle D_{jA}\bar{q}_{jz_j}, q_{jz_j}\rangle\, dz_{jA} &= \mathrm{Im}\langle D_{jA}\bar{q}_{jz_j}, D_{jB}q_{jz_j}\rangle\, dz_{jA} \wedge dz_{jB} \\ &= 2\,\mathrm{Im}\langle D_{jR}\bar{q}_{jz_j}, D_{jI}q_{jz_j}\rangle\, dz_{jR} \wedge dz_{jI} \\ &= 2\gamma(|z_j|^2)\, dz_{jR} \wedge dz_{jI} \\ &= \mathrm{i}\gamma_j(|z_j|^2)\, dz_j \wedge d\bar{z}_j, \end{aligned}$$

where $q_{jz_j} = z_j\hat{q}_j(|z_j|^2)$.

Turning to the proof of (4-14), we have

$$\Phi^*(\mathrm{i}\gamma_j(|z_j|^2)\, dz_j \wedge d\bar{z}_j) = \mathrm{i}\gamma_j(|z_j|^2)\, d(\Phi^* z_j) \wedge d(\Phi^*\bar{z}_j) = \mathrm{i}\gamma_j(|z_j|^2)\, dz_j \wedge d\bar{z}_j. \qquad \square$$

Lemma 4.4. *We have* $dB = \Omega$ *with* B *the differential form in the manifold* H^1 *defined by*

$$B(u)X := \mathrm{Im}\langle \bar{u}, X\rangle. \tag{4-15}$$

Consider, for $u \in B_{H^1}(0, d_0)$ *with the* d_0 *of Lemma 2.3, the function* $\psi \in C^\infty(B_{H^1}(0, d_0), \mathbb{R})$ *and the differential form* $\Gamma(u)$ *defined by*

$$\psi(u) := \sum_{j=1}^{n} \mathrm{Im}\langle \bar{q}_{jz_j}, u\rangle + \sum_{j=1}^{n} \mathrm{Im}(\alpha_j[z]\eta\bar{z}_j), \tag{4-16}$$

$$\Gamma(u) := B(u) - B_0(u) + d\psi(u). \tag{4-17}$$

Then the map $(z, \eta) \mapsto \Gamma(u(z, \eta))$, *where* $u(z, \eta)$ *is the right-hand side of* (2-10), *which is initially defined in* $B_{\mathbb{C}^n}(0, d_0) \times (H^1 \cap \mathcal{H}_c[0])$, *extends to* $B_{\mathbb{C}^n}(0, d_0) \times \Sigma^c_{-r}$ *for any* $r \in \mathbb{N}$. *In particular, we have* $\Gamma = \Gamma_{jA}dz_{jA} + \langle \Gamma_\eta, d\eta\rangle + \langle \Gamma_{\bar{\eta}}, d\bar{\eta}\rangle$ *with, in the sense of Remark 2.10,*

$$\Gamma_{jA} = \mathcal{R}^{1,1}_{\infty,\infty}(z, \mathbf{Z}, \eta) \qquad \text{and} \qquad \Gamma_\xi = \mathbf{S}^{1,1}_{\infty,\infty}(z, \mathbf{Z}, \eta) \quad \text{for } \xi = \eta, \bar{\eta}. \tag{4-18}$$

Furthermore, Γ *satisfies an invariance property in* $B_{H^1}(0, d_0)$:

$$\Phi^*\Gamma = \Gamma \quad \text{for } \Phi(u) = e^{\mathrm{i}\vartheta}u \text{ for any fixed } \vartheta \in \mathbb{R}. \tag{4-19}$$

Proof. By the definition of the exterior differential, and focusing on constant vector fields X and Y,

$$dB(X, Y) = XB(u)Y - YB(u)X = \mathrm{Im}\langle \overline{X}, Y\rangle - \mathrm{Im}\langle \overline{Y}, X\rangle = \Omega(X, Y).$$

This is enough to prove $dB = \Omega$. Next, using $R[z]\eta = \eta + \sum_j \alpha_j[z]\eta\phi_j$, we expand

$$B(u) = \sum_j \operatorname{Im}\langle \bar{Q}_{jz_j}, \cdot\rangle + \operatorname{Im}\langle \overline{R[z]\eta}, \cdot\rangle = \sum_j \operatorname{Im}\langle \bar{z}_j\phi_j, \cdot\rangle + \operatorname{Im}\langle\bar{\eta}, \cdot\rangle + \sum_j \operatorname{Im}\langle \bar{q}_{jz_j}, \cdot\rangle + \sum_j \operatorname{Im}(\overline{\alpha_j[z]\eta}\langle\phi_j, \cdot\rangle). \tag{4-20}$$

By the definition of B_0 in (4-13), we have

$$B - B_0 = I_1 + I_2 + I_3 + \sum_{j,A} \operatorname{Im}\langle D_{jA}\bar{q}_{jz_j}, q_{jz_j}\rangle dz_{jA} + \sum_j \operatorname{Im}\langle \bar{q}_{jz_j}, \cdot\rangle, \tag{4-21}$$

where

$$I_1 := \sum_j \operatorname{Im}[\bar{z}_j(\langle\phi_j, \cdot\rangle - dz_j)], \quad I_2 := -\operatorname{Im}\langle\bar{\eta}, d\eta - P_c\rangle, \quad I_3 := \sum_j \operatorname{Im}[\overline{\alpha_j[z]\eta}\langle\phi_j, \cdot\rangle].$$

We replace $d\eta$ using (4-4) and $\langle\phi_j, \cdot\rangle$ using (4-5). For $\alpha_j[z] \circ d\eta$, the linear operator defined by $\alpha_j[z] \circ d\eta(X) := \alpha_j[z]\, d\eta(X)$, we then get

$$I_1 = \operatorname{Im}\langle D_{jA}q_{jz_j}, \bar{z}_k\phi_k\rangle\, dz_{jA} + \operatorname{Im}(\bar{z}_j D_{kA}\alpha_j[z]\eta)\, dz_{kA} + \operatorname{Im}(\bar{z}_j\alpha_j[z] \circ d\eta) = \sum_{jA} \mathcal{R}^{1,1}_{\infty,\infty}\, dz_{jA} + \operatorname{Im}(\bar{z}_j\alpha_j[z] \circ d\eta), \tag{4-22}$$

where, as anticipated in Remark 2.10, here we set $\mathcal{R}^{i,j}_{K,M} = \mathcal{R}^{i,j}_{K,M}(z, \boldsymbol{Z}, \eta)$ and $\boldsymbol{S}^{i,j}_{K,M} = \boldsymbol{S}^{i,j}_{K,M}(z, \boldsymbol{Z}, \eta)$, where $\boldsymbol{Z}$ is as defined in Definition 2.2.

The second term in the last line of (4-22) is incorporated into the first sum in (4-25). We have

$$I_2 = \operatorname{Im}\langle\bar{\eta}, D_{jA}q_{jz_j}\rangle dz_{jA} = \sum_{jA} \mathcal{R}^{2,1}_{\infty,\infty} dz_{jA}. \tag{4-23}$$

Substituting with (4-5), we have

$$I_3 = \sum_{jA} \mathcal{R}^{2,1}_{\infty,\infty} dz_{jA} + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\eta\rangle + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\bar{\eta}\rangle. \tag{4-24}$$

Hence, we get

$$B - B_0 = \sum_j \operatorname{Im}(\bar{z}_j\alpha_j[z] \circ d\eta) + \sum_{jA} \mathcal{R}^{1,1}_{\infty,\infty}\, dz_{jA} + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\eta\rangle + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\bar{\eta}\rangle + \sum_{jA} \operatorname{Im}\langle D_{jA}\bar{q}_{jz_j}, q_{jz_j}\rangle\, dz_{jA} + \sum_j \operatorname{Im}\langle \bar{q}_{jz_j}, \cdot\rangle. \tag{4-25}$$

Set now $\widetilde{\psi}(u) := -\sum_{j=1}^n \operatorname{Im}\langle \bar{q}_{jz_j}, u\rangle$. Then it is elementary that we have

$$d\widetilde{\psi} = -\sum_{j=1}^n \operatorname{Im}\langle \bar{q}_{jz_j}, \cdot\rangle - \sum_{j,A} \operatorname{Im}\langle D_{jA}\bar{q}_{jz_j}, q_{jz_j}\rangle\, dz_{jA} + \sum_{j,A} \mathcal{R}^{1,1}_{\infty,\infty}\, dz_{jA}. \tag{4-26}$$

By the Leibniz rule we have

$$\operatorname{Im}(\bar{z}_j\alpha_j[z] \circ d\eta) = d\operatorname{Im}(\bar{z}_j\alpha_j[z]\, \eta) - \operatorname{Im}(d(\bar{z}_j\alpha_j[z])\, \eta). \tag{4-27}$$

The contribution to $\sum_j \operatorname{Im}(\bar z_j \alpha_j[z] \circ d\eta)$ in (4-25) of the last term in the right-hand side of (4-27) can be absorbed into the term $\sum_{jA} \mathcal{R}^{1,1}_{\infty,\infty} dz_{jA}$. Then

$$B - B_0 + d\psi = \sum_{jA} \mathcal{R}^{1,1}_{\infty,\infty} dz_{jA} + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\eta\rangle + \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\bar\eta\rangle.$$

Here we have used that the first two terms in the right-hand side of (4-26) cancel with the last two sums in (4-25) and that there is a cancellation between the contribution to $\sum_j \operatorname{Im}(\bar z_j \alpha_j[z] \circ d\eta)$ of the $d \operatorname{Im}(\bar z_j \alpha_j[z]\, \eta)$ in (4-27) and the differential of the last term in (4-16). This yields (4-18).

Lastly we consider (4-19). We have $\Phi^* B_0 = B_0$ by (4-14), while $\Phi^* B = B$ follows immediately from the definition of B in (4-15). Finally, $\Phi^*\psi = \psi$ follows immediately from $\Phi^*\langle \bar q_{jz_j}, u\rangle = \langle \bar q_{jz_j}, u\rangle$, which follows from $q_{jz_j}(e^{\mathrm{i}\vartheta} z) = e^{\mathrm{i}\vartheta} q_{jz_j}(z)$, and from (2-9) and (2-12), which imply

$$\Phi^*(\bar z_j \alpha_j[z]\eta) = e^{-\mathrm{i}\vartheta} \bar z_j \alpha_j[e^{\mathrm{i}\vartheta} z] e^{\mathrm{i}\vartheta}\eta = \bar z_j \alpha_j[z]\eta. \qquad \square$$

Lemma 4.5. *Consider the differential form $\Omega - \Omega_0$, which is defined in $B_{H^1}(0, d_0)$ for the d_0 of Lemma 2.3. Then, summing on repeated indexes, we have*

$$\Omega - \Omega_0 = \widetilde\Omega_{ijAB}\, dz_{iA} \wedge dz_{jB} + \sum_{\xi=\eta,\bar\eta} dz_{iA} \wedge \langle \widetilde\Omega_{iA\xi}, d\xi\rangle + \sum_{\xi,\xi'=\eta,\bar\eta} \langle \widetilde\Omega_{\xi'\xi}\, d\xi, d\xi'\rangle, \tag{4-28}$$

where, expressed as functions of (z,η), the coefficients extend into functions defined in $B_{\mathbb{C}^n}(0,d_0) \times \Sigma^c_{-r}$ for any $r \in \mathbb{N}$ and, in particular, we have $\widehat\Omega_{iA\xi} = \boldsymbol{S}^{1,0}_{\infty,\infty}(z, \boldsymbol{Z}, \eta)$, $\widehat\Omega_{ijAB} = \mathcal{R}^{1,0}_{\infty,\infty}(z, \boldsymbol{Z}, \eta)$ in the sense of Remark 2.10 and $\widetilde\Omega_{\xi'\xi} = \partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}(z, \boldsymbol{Z}, \eta) - (\partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}(z, \boldsymbol{Z}, \eta))^*$ (with the two instances of $\boldsymbol{S}$ distinct). We furthermore have

$$\Phi^*(\Omega - \Omega_0) = \Omega - \Omega_0 \quad \text{for } \Phi(z,\eta) = (e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) \text{ for any fixed } \vartheta \in \mathbb{R}. \tag{4-29}$$

Proof. We have

$$\Omega - \Omega_0 = d\Gamma = d \sum_{j,A} \mathcal{R}^{1,1}_{\infty,\infty} dz_{jA} + d \sum_{\xi} \langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\xi\rangle.$$

Summing over k, B and ξ, we have

$$d(\mathcal{R}^{1,1}_{\infty,\infty}\, dz_{jA}) = \partial_{z_{kB}} \mathcal{R}^{1,1}_{\infty,\infty}\, dz_{kB} \wedge dz_{jA} + \langle \partial_\xi \mathcal{R}^{1,1}_{\infty,\infty}, d\xi\rangle \wedge dz_{jA}$$

with the $\partial_\xi \mathcal{R}^{1,1}_{\infty,\infty} \in \mathcal{H}_c[0]$ defined, summing on repeated indexes and for F with values in $\mathbb{R}$, by

$$dF X = \partial_{z_{kB}} F\, dz_{kB}\, X + \langle \partial_\xi F, d\xi\, X\rangle \quad \text{for any } X \in L^2(\mathbb{R}^3, \mathbb{C}).$$

It is easy to see that $\partial_\xi \mathcal{R}^{1,1}_{\infty,\infty} = \boldsymbol{S}^{1,0}_{\infty,\infty}$ and $\partial_{z_{kB}} \mathcal{R}^{1,1}_{\infty,\infty} = \mathcal{R}^{1,0}_{\infty,\infty}$.

Furthermore, summing on repeated indexes we have

$$\begin{aligned} d\langle \boldsymbol{S}^{1,1}_{\infty,\infty}, d\xi\rangle &= dz_{kB} \wedge \langle \partial_{z_{kB}} \boldsymbol{S}^{1,1}_{\infty,\infty}, d\xi\rangle + \langle \partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}\, d\xi', d\xi\rangle - \langle d\xi, \partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}\, d\xi'\rangle \\ &= dz_{kB} \wedge \langle \partial_{z_{kB}} \boldsymbol{S}^{1,1}_{\infty,\infty}, d\xi\rangle + \langle \partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}\, d\xi', d\xi\rangle - \langle (\partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty})^*\, d\xi, d\xi'\rangle, \end{aligned} \tag{4-30}$$

where, for $T \in C^1(U_{L^2}, L^2)$ with U_{L^2} an open subset in L^2, $\partial_\xi T \in B(\mathcal{H}_c[0], L^2)$ is defined by

$$dT\,X = \partial_{z_{kB}} T\, dz_{kB}\, X + \partial_\xi T\, d\xi\, X \quad \text{for any } X \in L^2(\mathbb{R}^3, \mathbb{C}).$$

Summing on ξ in (4-30) we get terms which are absorbed into the last two terms of (4-28).

Formula (4-29) follows from (4-19), $\Omega_0 = dB_0$ and $\Omega = dB$. □

Lemma 4.6. *Consider the form $\Omega_t := \Omega_0 + t(\Omega - \Omega_0)$ and set $i_X \Omega_t(Y) := \Omega_t(X, Y)$. For any preassigned $r \in \mathbb{N}$ recall by, (4-8), (4-13) and Lemmas 4.4 and 4.5, that $\Omega - \Omega_0$ and Γ extend to forms defined in $B_{\mathbb{C}^n}(0, d_0) \times \Sigma^c_{-r}$. Then there is $\delta_0 \in (0, d_0)$ such that, for any $(t, z, \eta) \in (-4, 4) \times B_{\mathbb{C}^n}(0, \delta_0) \times B_{\Sigma^c_{-r}}(0, \delta_0)$, there exists exactly one solution $\mathcal{X}^t(z, \eta) \in L^2$ of the equation $i_{\mathcal{X}^t} \Omega_t = -\Gamma$. Furthermore, we have the following facts*:

(1) *$\mathcal{X}^t(z, \eta) \in \Sigma_r$ and, if we set $\mathcal{X}^t_{jA}(z, \eta) = dz_{jA}\, \mathcal{X}^t(z, \eta)$ and $\mathcal{X}^t_\eta(z, \eta) = d\eta\, \mathcal{X}^t(z, \eta)$, we have $\mathcal{X}^t_{jA}(z, \eta) = \mathcal{R}^{1,1}_{r,\infty}(t, z, \boldsymbol{Z}, \eta)$ and $\mathcal{X}^t_\eta(z, \eta) = \boldsymbol{S}^{1,1}_{r,\infty}(t, z, \boldsymbol{Z}, \eta)$ in the sense of Remark 2.10.*

(2) *For $\mathcal{X}^t_j := dz_j\, \mathcal{X}^t$ and $\mathcal{X}^t_\eta := d\eta\, \mathcal{X}^t$, we have $\mathcal{X}^t_j(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} \mathcal{X}^t_j(z, \eta)$ and $\mathcal{X}^t_\eta(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} \mathcal{X}^t_\eta(z, \eta)$.*

Proof. We define Y such that $i_Y \Omega'_0 = -\Gamma$, which yields $Y_{jR} = -\frac{1}{2}\Gamma_{jI}$ and $Y_{jI} = \frac{1}{2}\Gamma_{jR}$ (both $\mathcal{R}^{1,1}_{\infty,\infty}$), $Y_\eta = -\mathrm{i}\Gamma_{\bar\eta}$ and $Y_{\bar\eta} = \mathrm{i}\Gamma_\eta$ (both $\boldsymbol{S}^{1,1}_{\infty,\infty}$). We use $i_{K_t X}\Omega'_0 = i_X(\Omega_0 - \Omega'_0 + t\widehat{\Omega})$, where $\widehat{\Omega} := \Omega - \Omega_0$, to define in L^2 the operator K_t. We claim the following lemma:

Lemma 4.7. *For appropriate symbols $\mathcal{R}^{1,0}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta)$ and $\boldsymbol{S}^{1,0}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta)$, which differ from one term to the other, and for $\boldsymbol{Z}$ as in Definition 2.2, we have*

$$\begin{aligned}(K_t X)_{jA} &= \sum_{lB} \mathcal{R}^{1,0}_{\infty,\infty} X_{lB} + \sum_{\xi=\eta,\bar\eta} \langle \boldsymbol{S}^{1,0}_{\infty,\infty}, X_\xi\rangle,\\ (K_t X)_\xi &= \sum_{lB} \boldsymbol{S}^{1,0}_{\infty,\infty} X_{lB} + \sum_{\xi'=\eta,\bar\eta} \big(\partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta) - (\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta))^*\big) X_{\xi'}.\end{aligned} \tag{4-31}$$

We assume for a moment Lemma 4.7 and complete the proof of Lemma 4.6. The equation $i_{\mathcal{X}^t}\Omega_t = -\Gamma$ becomes $\mathcal{X}^t + K_t \mathcal{X}^t = Y$. Indeed, suppose $\mathcal{X}^t + K_t\mathcal{X}^t = Y$ holds. Then, by definition of K_t, we have

$$i_{\mathcal{X}_t}(\Omega_t - \Omega') = i_{K_t \mathcal{X}_t}\Omega'_0 \quad \text{and so} \quad i_{\mathcal{X}_t}\Omega_t = i_{\mathcal{X}_t}\Omega'_0 + i_{K_t\mathcal{X}_t}\Omega'_0 = -\Gamma.$$

By Lemma 4.7, in coordinates and for $\xi = \eta, \bar\eta$, the last equation is schematically of the form

$$\begin{aligned}\mathcal{X}^t_{jA} + \sum_{\ell, B} \mathcal{R}^{1,0}_{r,\infty} \mathcal{X}^t_{\ell B} + \sum_{\xi=\eta,\bar\eta} \langle \boldsymbol{S}^{1,1}_{r,\infty}, \mathcal{X}^t_\xi\rangle &= \mathcal{R}^{1,1}_{r,\infty},\\ \mathcal{X}^t_\xi + \sum_{\ell B} \boldsymbol{S}^{1,0}_{r,\infty}\mathcal{X}^t_{\ell B} + \sum_{\xi'=\eta,\bar\eta} \big(\partial_{\xi'} \boldsymbol{S}^{1,1}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta) - (\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}(t, z, \boldsymbol{Z}, \eta))^*\big)\mathcal{X}^t_{\xi'} &= \boldsymbol{S}^{1,1}_{r,\infty}.\end{aligned} \tag{4-32}$$

Notice that $(\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty})\boldsymbol{S}^{1,1}_{r,\infty}$ is C^∞ in $(t, z, \boldsymbol{Z}, \eta)$ with values in Σ_r. We have

$$\|(\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty})\boldsymbol{S}^{1,1}_{r,\infty}\|_{\Sigma_r} \le \|\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}\|_{B(\Sigma_{-r},\Sigma_r)} \|\boldsymbol{S}^{1,1}_{r,\infty}\|_{\Sigma_r}.$$

By (2-26), we have $\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}(t,0,0,0)$. This implies

$$\|\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}\|_{B(\Sigma_{-r},\Sigma_r)} \le C\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}| + |z| \tag{4-33}$$

and so

$$\|(\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty})\boldsymbol{S}^{1,1}_{r,\infty}\|_{\Sigma_r} \le C(\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}|)(\|\eta\|_{\Sigma_{-K}} + |\boldsymbol{Z}| + |z|)^2.$$

So $(\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty})\boldsymbol{S}^{1,1}_{r,\infty} = \boldsymbol{S}^{2,1}_{r,\infty}$.

Inequality (4-33), a Neumann expansion and formulas (2-27) yield claim (1) in Lemma 4.6.

Claim (2) in Lemma 4.6 follows from

$$i_{\Phi_*^{-1}\mathscr{X}^t}\Phi^*\Omega_t = -\Phi^*\Gamma = -\Gamma = i_{\mathscr{X}^t}\Omega_t = i_{\Phi_*^{-1}\mathscr{X}^t}\Omega_t,$$

where $\Phi^*\Gamma = \Gamma$ is (4-19) and we use (4-14) and (4-29) to conclude $\Phi^*\Omega_t = \Omega_t$. Then $\Phi_*^{-1}\mathscr{X}^t = \mathscr{X}^t$, which is equivalent to $\Phi_*\mathscr{X}^t = \mathscr{X}^t$. For the other formulas in claim (2), we have, for instance,

$$\mathscr{X}^t_j(e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta) = \mathscr{X}^t_j(\Phi(u)) = dz_j\,(\mathscr{X}^t(\Phi(u))) = dz_j(\Phi_*\mathscr{X}^t(u)) = d(z_j\circ\Phi)(\mathscr{X}^t(u)) = e^{\mathrm{i}\vartheta}\mathscr{X}^t_j(u).$$

This ends the proof of Lemma 4.6, assuming Lemma 4.7. □

Proof of Lemma 4.7. By (4-13) and summing over the indexes (j, A, B), we can write

$$\Omega_0 - \Omega_0' = \mathscr{R}^{4,0}_{\infty,\infty}\,dz_{jA}\wedge dz_{jB} \implies i_X(\Omega_0 - \Omega_0') = \mathscr{R}^{4,0}_{\infty,\infty}X_{jR}\,dz_{jI} + \mathscr{R}^{4,0}_{\infty,\infty}X_{jI}\,dz_{jR}. \tag{4-34}$$

So, if we define $K'X$ by setting $i_{K'X}\Omega_0' = i_X(\Omega_0 - \Omega_0')$, by comparing (4-34) with

$$i_{K'X}\Omega_0' = 2(K'X)_{jR}dz_{jI} - 2(K'X)_{jI}dz_{jR} + \mathrm{i}\langle (K'X)_\eta, X_{\bar\eta}\rangle - \mathrm{i}\langle (K'X)_{\bar\eta}, X_\eta\rangle,$$

we obtain

$$(K'X)_{jA} = \mathscr{R}^{4,0}_{\infty,\infty}X_{jA} \qquad \text{and} \qquad (K'X)_\xi = 0 \quad \text{for } \xi = \eta, \bar\eta. \tag{4-35}$$

Summing on (j, l, A, B, ξ, ξ'), we have

$$t\widehat{\Omega} = \mathscr{R}^{1,0}_{\infty,\infty}\,dz_{jA}\wedge dz_{lB} + dz_{jA}\wedge\langle \boldsymbol{S}^{1,0}_{\infty,\infty}, d\xi\rangle + t\big\langle\big(\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty}(z,\boldsymbol{Z},\eta) - (\partial_{\xi'}\boldsymbol{S}^{1,1}_{\infty,\infty}(z,\boldsymbol{Z},\eta))^*\big)\,d\xi, d\xi'\big\rangle.$$

Hence,

$$ti_X\widehat{\Omega} = \mathscr{R}^{1,0}_{\infty,\infty}X_{jA}\,dz_{lB} + \langle \boldsymbol{S}^{1,0}_{\infty,\infty}, X_\xi\rangle\,dz_{jA} + X_{jA}\langle \boldsymbol{S}^{1,0}_{\infty,\infty}, d\xi\rangle + \langle[\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty} - (\partial_{\xi'}\boldsymbol{S}^{1,1}_{\infty,\infty})^*]X_\xi, d\xi'\rangle.$$

So, if we define $K''X$ by setting $i_{K''X}\Omega_0' = ti_X\widehat{\Omega}$, we obtain

$$\begin{aligned}(K''X)_{jA} &= \sum_{\ell B}\mathscr{R}^{1,0}_{\infty,\infty}X_{\ell B} + \sum_{\xi=\eta,\bar\eta}\langle \boldsymbol{S}^{1,0}_{\infty,\infty}, X_\xi\rangle,\\ (K''X)_\xi &= \sum_{lB}\boldsymbol{S}^{1,0}_{\infty,\infty}X_{lB} + \sum_{\xi=\eta,\bar\eta}[\partial_{\xi'}\boldsymbol{S}^{1,1}_{\infty,\infty} - (\partial_\xi \boldsymbol{S}^{1,1}_{\infty,\infty})^*]X_{\xi'}.\end{aligned} \tag{4-36}$$

Since $K_t = K' + K''$, summing up (4-35) and (4-36) we get (4-31), and so Lemma 4.7. □

Having established that $\mathscr{X}^t(z,\eta)$ has components which are restrictions of symbols as in Definitions 2.8 and 2.9, we have the following result:

Lemma 4.8. *Fix $r \in \mathbb{N}$ and for the δ_0 and the $\mathcal{X}^t(z,\eta)$ of Lemma 4.6, consider the following system, which is well defined in $(t,z,\eta) \in (-4,4) \times B_{\mathbb{C}^n}(0,\delta_0) \times B_{\Sigma_k^c}(0,\delta_0)$ for all $k \in \mathbb{Z} \cap [-r,r]$:*

$$\dot{z}_j = \mathcal{X}^t_j(z,\eta) \quad \text{and} \quad \dot{\eta} = \mathcal{X}^t_\eta(z,\eta). \tag{4-37}$$

Then the following facts hold:

(1) *For $\delta_1 \in (0,\delta_0)$ sufficiently small, system (4-37) generates flows, for all $k \in \mathbb{Z} \cap [-r,r]$,*

$$\begin{aligned} &\mathfrak{F}^t \in C^\infty((-2,2) \times B_{\mathbb{C}^n}(0,\delta_1) \times B_{\Sigma_k^c}(0,\delta_1), B_{\mathbb{C}^n}(0,\delta_0) \times B_{\Sigma_k^c}(0,\delta_0)),\\ &\mathfrak{F}^t \in C^\infty((-2,2) \times B_{\mathbb{C}^n}(0,\delta_1) \times B_{H^1\cap\mathcal{H}_c[0]}(0,\delta_1), B_{\mathbb{C}^n}(0,\delta_0) \times B_{H^1\cap\mathcal{H}_c[0]}(0,\delta_0)).\end{aligned} \tag{4-38}$$

In particular, for $z_j^t := z_j \circ \mathfrak{F}^t(z,\eta)$ and $\eta^t := \eta \circ \mathfrak{F}^t(z,\eta)$, we have

$$z_j^t = z_j + S_j(t,z,\eta) \quad \text{and} \quad \eta^t = \eta + S_\eta(t,z,\eta) \tag{4-39}$$

with $S_j(t,z,\eta) = \mathcal{R}^{1,1}_{r,\infty}(t,z,\boldsymbol{Z},\eta)$ and $S_\eta(t,z,\eta) = \boldsymbol{S}^{1,1}_{r,\infty}(t,z,\boldsymbol{Z},\eta)$ in the sense of Remark 2.10.

(2) *$\mathfrak{F} = \mathfrak{F}^1$ is a local diffeomorphism of H^1 into itself near the origin such that $\mathfrak{F}^*\Omega = \Omega_0$.*

(3) *$S_j(t, e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} S_j(t,z,\eta)$ and $S_\eta(t, e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} S_\eta(t,z,\eta)$.*

Proof. The first sentence has been established in Lemma 4.6. Elementary theory of ODEs yields (4-38). The rest of claim (1) is a special case of a more general result; see Lemma 4.9 below. We get claim (2) by the classical formula, for L_X the Lie derivative,

$$\partial_t(\mathfrak{F}^{t*}\Omega_t) = \mathfrak{F}^{t*}(L_{\mathcal{X}^t}\Omega_t + \partial_t\Omega_t) = \mathfrak{F}^{t*}(di_{\mathcal{X}^t}\Omega_t + d\Gamma) = 0. \tag{4-40}$$

Notice that (4-40) is well defined here, while it has no clear meaning for the NLS with translation treated in [Cuccagna 2012; 2014], where the flows $\mathfrak{F}^t$ are not differentiable (see [Cuccagna 2012] for a rigorous argument on how to get around this problem). The symmetry in claim (3) is elementary and we skip it. ☐

Lemma 4.9. *Consider a system*

$$\dot{z}_j = X_j(t,z,\eta) \quad \text{and} \quad \dot{\eta} = X_\eta(t,z,\eta), \tag{4-41}$$

where $X_j = \mathcal{R}^{a,b}_{r,m}(t,z,\boldsymbol{Z},\eta)$ for all j and $X_\eta = \boldsymbol{S}^{c,d}_{r,m}(t,z,\boldsymbol{Z},\eta)$ for fixed pairs (r,m), (a,b) and (c,d). Assume $m,b,d \geq 1$, with possibly $m=\infty$, and $r \geq 0$. Then, for the flow $(z^t,\eta^t) = \mathfrak{F}^t(z,\eta)$, we have

$$z_j^t = z_j + S_j(t,z,\eta) \quad \text{and} \quad \eta^t = \eta + S_\eta(t,z,\eta) \tag{4-42}$$

for appropriate functions $S_j = \mathcal{R}^{a,b}_{r,m}(t,z,\boldsymbol{Z},\eta)$ and $S_\eta = \boldsymbol{S}^{c,d}_{r,m}(t,z,\boldsymbol{Z},\eta)$ in the sense of Remark 2.10.

Proof. Consider the vectors $\boldsymbol{Z}$ of Definition 2.2. Notice that $\dot{\boldsymbol{Z}} = \mathcal{R}^{a+1,b}_{r,m}(t,z,\boldsymbol{Z},\eta)$, and this equation can be extended to a whole neighborhood of 0 in the space L. Pairing the latter equation with equations (4-42), a system remains defined which has a flow $\mathfrak{F}^t(z,\boldsymbol{Z},\eta)$ that is C^m in $(t,z,\boldsymbol{Z},\eta)$ and which reduces to the flow in (4-41) when we restrict to the vectors $\boldsymbol{Z}$ of Definition 2.2, by construction. The inequalities

(2-23) and (2-26), required to prove $S_j = \mathcal{R}^{a,b}_{r,m}$ and $S_\eta = \boldsymbol{S}^{c,d}_{r,m}$, can be obtained as follows. We have, for all $|k| \le r$,

$$
\begin{aligned}
|z^t - z| &\le \int_0^t |\mathcal{R}^{a,b}_{r,m}(s, z^s, \boldsymbol{Z}^s, \eta^s)|\, ds \\
&\le C \int_0^t (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s|)^b (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s| + |z^s|)^a\, ds, \\
\|\eta^t - \eta\|_{\Sigma_k} &\le \int_0^t \|\boldsymbol{S}^{c,d}_{r,m}(s, z^s, \boldsymbol{Z}^s, \eta^s)\|_{\Sigma_k}\, ds \\
&\le C \int_0^t (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s|)^d (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s| + |z^s|)^c\, ds, \\
|\boldsymbol{Z}^t - \boldsymbol{Z}| &\le \int_0^t |\mathcal{R}^{a,b}_{r,m}(s, z^s, \boldsymbol{Z}^s, \eta^s)|\, ds \\
&\le C \int_0^t (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s|)^b (\|\eta^s\|_{\Sigma_{-r}} + |\boldsymbol{Z}^s| + |z^s|)^{a+1}\, ds.
\end{aligned} \tag{4-43}
$$

By Gronwall's inequality we get that $|\boldsymbol{Z}^t|$ and $\|\eta^t\|_{\Sigma_{-r}}$ are bounded by $C(|\boldsymbol{Z}| + \|\eta\|_{\Sigma_{-r}})$. Plugging this into the right-hand side of (4-43), we obtain the last part of the statement. □

We discuss the pullback of the energy E by the map $\mathfrak{F} := \mathfrak{F}^1$ in Lemma 4.8(2). We set $H_2(z, \eta) = \sum_{j=1}^n e_j |z_j|^2 + \langle H\eta, \bar\eta\rangle$. Our first preliminary result is the following one:

Lemma 4.10. *Consider the δ_1 of Lemma 4.8, the δ_0 of Lemma 4.6 and set $r = r_0$ with r_0 the index in Lemma 3.1. Then, for the map $\mathfrak{F}$ in Lemma 4.8(2), we have*

$$
\mathfrak{F}\big(B_{\mathbb{C}^n}(0, \delta_1) \times (B_{H^1}(0, \delta_1) \cap \mathcal{H}_c[0])\big) \subset B_{\mathbb{C}^n}(0, \delta_0) \times (B_{H^1}(0, \delta_0) \cap \mathcal{H}_c[0]) \tag{4-44}
$$

and $\mathfrak{F}|_{B_{\mathbb{C}^n}(0,\delta_1)\times(B_{H^1}(0,\delta_1)\cap\mathcal{H}_c[0])}$ is a diffeomorphism between its domain and an open neighborhood of the origin in $\mathbb{C}^n \times (H^1 \cap \mathcal{H}_c[0])$. Furthermore, the functional $K := E \circ \mathfrak{F}$ admits an expansion

$$
\begin{aligned}
&K(z, \eta) \\
&= H_2(z, \eta) + \sum_{j=1,\dots,n} \lambda_j(|z_j|^2) \\
&\quad + \sum_{l=0}^{2N+3} \sum_{|\boldsymbol{m}|=l+1} \boldsymbol{Z}^{\boldsymbol{m}} a^{(1)}_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) + \sum_{j=1}^{n} \sum_{l=0}^{2N+3} \sum_{|\boldsymbol{m}|=l} (\bar z_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G^{(1)}_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}) \\
&\qquad + \mathcal{R}^{1,2}_{r_1,\infty}(z, \eta) + \mathcal{R}^{0,2N+5}_{r_1,\infty}(z, \boldsymbol{Z}, \eta) + \operatorname{Re}\langle \boldsymbol{S}^{0,2N+4}_{r_1,\infty}(z, \boldsymbol{Z}, \eta), \bar\eta\rangle \\
&\qquad\quad + \sum_{i+j=2} \sum_{|\boldsymbol{m}|\le 1} \boldsymbol{Z}^{\boldsymbol{m}} \langle G^{(1)}_{2\boldsymbol{m}ij}(z, \eta), \eta^i \bar\eta^j\rangle + \sum_{d+c=3} \sum_{i+j=d} \langle G^{(1)}_{dij}(z), \eta^i \bar\eta^j\rangle \mathcal{R}^{0,c}_{r,\infty}(z, \eta) + E_P(\eta),
\end{aligned} \tag{4-45}
$$

where $r_1 = r_0 - 2$, $G^{(1)}_{j\boldsymbol{m}}$, $G^{(1)}_{2\boldsymbol{m}ij}$ and $G^{(1)}_{dij}$ are $\boldsymbol{S}^{0,0}_{r_1,\infty}$, $a^{(1)}_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) = \mathcal{R}^{0,0}_{\infty,\infty}(z)$, c.c. means complex conjugate, and $\lambda_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}(|z_j|^2)$. For $|\boldsymbol{m}| = 0$, $G^{(1)}_{2\boldsymbol{m}ij}(z, \eta) = G_{2\boldsymbol{m}ij}(z)$ is the same as (3-4). Finally, we have the invariance $\mathcal{R}^{1,2}_{r_1,\infty}(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r_1,\infty}(z, \eta)$.

Proof. Consider the expansion (3-3) for $E(u(z',\eta'))$, and substitute the formulas $z'_j = z_j + S_j(z,\eta)$ and $\eta' = \eta + S_\eta(z,\eta)$, with $S_\ell(z,\eta) = S_\ell(1,z,\eta)$ for $\ell = j, \bar{j}, \eta, \bar{\eta}$, with $S_{\bar{\ell}} = \bar{S}_\ell$. By $S_j(z,\eta) = \mathcal{R}^{1,1}_{r_0,\infty}(z,\boldsymbol{Z},\eta)$ and $S_\eta(z,\eta) = \boldsymbol{S}^{1,1}_{r_0,\infty}(z,\boldsymbol{Z},\eta)$, it is elementary to see that the last three lines of (3-3) yield terms that can be absorbed into the last three lines (4-45) (with $l \ge 1$ in the third line). Notice that the z dependence of the $a^{(1)}_{\boldsymbol{m}}$ in terms of $(|z_1|^2, \dots, |z_n|^2)$ follows by Lemmas 4.8 and B.3. The z dependence of the $G^{(1)}_{j\boldsymbol{m}}$ is obtained by Lemma B.4. Notice also that, if an $\mathcal{R}^{i,0}_{r,\infty}(z)$ depends only on z, then it is an $\mathcal{R}^{i,0}_{\infty,\infty}(z)$.

We have $\mathcal{R}^{1,2}_{r_0,\infty}(z',\eta') = \mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},\eta)$. Note that, by the invariance of $\mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)$ and Lemma 4.8(3), we have $\mathcal{R}^{1,2}_{r_0,\infty}(e^{\mathrm{i}\vartheta}z,\boldsymbol{Z},e^{\mathrm{i}\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},\eta)$. By Taylor expansion (using the conventions under (3-14))

$$\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},\eta) = \mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},0) + d_\eta\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},0)\eta + \int_0^1 (1-t)\partial_\eta^2\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},t\eta)\,dt\cdot\eta^2. \tag{4-46}$$

Each of the terms in the right-hand side is invariant by change of variables $(z,\eta) \rightsquigarrow (e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta)$. We have

$$\begin{aligned}\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},\eta)|_{\eta=0} = \mathcal{R}^{1,2}_{\infty,\infty}(z,\boldsymbol{Z}) &= \sum_{k\le 2N+4}\frac{1}{k!}d^k_{\boldsymbol{Z}}\mathcal{R}^{1,2}_{\infty,\infty}(z,0)\boldsymbol{Z}^k + \mathcal{R}^{1,2N+5}_{\infty,\infty}(z,\boldsymbol{Z})\\ &= \mathcal{R}^{1,2N+5}_{\infty,\infty}(z,\boldsymbol{Z}) + \sum_{l=2}^{2N+4}\sum_{|\boldsymbol{m}|=l+1}\boldsymbol{Z}^{\boldsymbol{m}}c_{\boldsymbol{m}}(z)\\ &= \mathcal{R}^{1,2N+5}_{\infty,\infty}(z,\boldsymbol{Z}) + \sum_{l=2}^{2N+4}\sum_{|\boldsymbol{m}|=l+1}\boldsymbol{Z}^{\boldsymbol{m}}\sum_{j=1}^{n}c_{j\boldsymbol{m}}(|z_j|^2),\end{aligned}$$

where, as in Step 1 in Lemma 3.1, the last equality is obtained by the invariance with respect to $(z,\eta) \rightsquigarrow (e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta)$ and by smoothness. We have, proceeding as above,

$$\begin{aligned}d_\eta\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},0)\eta &= \mathrm{Re}\langle \boldsymbol{S}^{1,1}_{r_0,\infty}(z,\boldsymbol{Z}),\bar{\eta}\rangle\\ &= \sum_{k\le 2N+3}\frac{1}{k!}\mathrm{Re}\langle d^k_{\boldsymbol{Z}}\boldsymbol{S}^{1,1}_{r_0,\infty}(z,0),\bar{\eta}\rangle\boldsymbol{Z}^k + \mathrm{Re}\langle \boldsymbol{S}^{1,2N+4}_{r_0,\infty}(z,\boldsymbol{Z},\eta),\bar{\eta}\rangle\\ &= \mathrm{Re}\langle \boldsymbol{S}^{1,2N+4}_{r_0,\infty}(z,\boldsymbol{Z},\eta),\bar{\eta}\rangle + \sum_{j=1}^{n}\sum_{l=1}^{2N+3}\sum_{|\boldsymbol{m}|=l}(\bar{z}_j\boldsymbol{Z}^{\boldsymbol{m}}\langle A_{j\boldsymbol{m}}(|z_j|^2),\eta\rangle + \text{c.c.}),\end{aligned}$$

Finally, for an $\mathcal{R}^{1,2}_{r_0,\infty}(e^{\mathrm{i}\vartheta}z, e^{\mathrm{i}\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r_0,\infty}(z,\eta)$ we have — see Definition 2.8 —

$$\int_0^1 (1-t)\partial_\eta^2\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},t\eta)\,dt\,\eta^2 = \mathcal{R}^{1,2}_{r_0,\infty}(z,\eta).$$

By (4-46) and the subsequent formulas, we see that $\mathcal{R}^{1,2}_{r_0,\infty}(z',\eta')$ is absorbed into the last three lines of (4-45) (with $l \ge 1$ in the third line). The term $\langle H\eta',\bar{\eta}'\rangle = \langle H\eta,\bar{\eta}\rangle + \mathcal{R}^{1,2}_{r_0-2,\infty}(z,\boldsymbol{Z},\eta)$ behaves similarly, recalling that $r_1 = r_0 - 2$. Here too we have $\mathcal{R}^{1,2}_{r_0-2,\infty}(e^{\mathrm{i}\vartheta}z,\boldsymbol{Z},e^{\mathrm{i}\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r_0-2,\infty}(z,\boldsymbol{Z},\eta)$. This function can be treated like the $\mathcal{R}^{1,2}_{r_0,\infty}(z,\boldsymbol{Z},\eta)$ discussed earlier.

The terms $E(Q_{jz_j})$ and, for $j \ne k$, $\mathrm{Re}\langle q_{jz_j}, \bar{z}_k\phi_k\rangle = \mathcal{R}^{1,1}_{\infty,\infty}(z,\boldsymbol{Z})$ can be expanded similarly. But this time we need $l = 0$ in the third line. $\square$

The expansion in Lemma 4.10 is too crude. We have the following additional and crucial fact:

Lemma 4.11 (cancellation lemma). *In the third line of (4-45) all the terms with $l = 0$ are zeros.*

Proof. We first observe that the terms in the third line of (4-45) with $l = 0$ can be written as

$$\sum_{k=1}^{n}\sum_{j\neq k}\sum_{A=R,I} z_{jA} b_{kjA}(z_k) + \sum_{k=1}^{n} \operatorname{Re}\langle \boldsymbol{A}_k(z_k), \bar{\eta}\rangle. \tag{4-47}$$

Indeed, they are

$$\sum_{|\boldsymbol{m}|=1} \boldsymbol{Z}^{\boldsymbol{m}} a^{(1)}_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) + \sum_{j=1}^{n} (\bar{z}_j \langle G^{(1)}_{j\boldsymbol{0}}(|z_j|^2), \eta\rangle + \text{c.c.}), \tag{4-48}$$

and it is obvious that the second term of (4-48) is the second term of (4-47). Arguing as in Lemma 3.1, the first term of (4-48) can be written as

$$\sum_{k=1}^{n}\sum_{|\boldsymbol{m}|=1} \boldsymbol{Z}^{\boldsymbol{m}} a^{(1)}_{k\boldsymbol{m}}(|z_k|^2).$$

Further, for $\boldsymbol{Z}^{\boldsymbol{m}} = z_i\bar{z}_j$, we can assume that i or j must be equal to k, because, if not, it can be absorbed into the terms with $l \geq 1$. Set $\mathcal{N}_k := \{\boldsymbol{m} : |\boldsymbol{m}| = 1,\ m_{i,j} = 0 \text{ if } i \neq k \text{ and } j \neq k\}$. We have

$$\sum_{k=1}^{n}\sum_{|\boldsymbol{m}|=1} \boldsymbol{Z}^{\boldsymbol{m}} a^{(1)}_{k\boldsymbol{m}}(|z_k|^2) = \sum_{k=1}^{n}\sum_{\boldsymbol{m}\in\mathcal{N}_k} \boldsymbol{Z}^{\boldsymbol{m}} a^{(1)}_{k\boldsymbol{m}}(|z_k|^2) = \sum_{k=1}^{n}\sum_{j\neq k} (z_j\bar{z}_k a^{(1)}_{km_{jk}}(|z_k|^2) + z_k\bar{z}_j a^{(1)}_{km_{jk}}(|z_k|^2)).$$

So, we can write the term in the form of the first term of (4-47).

Next, notice that, for $p_k = (0, \dots, 0, z_k, \dots, 0; 0)$,

$$b_{kjA}(z_k) = \partial_{z_{jA}} K(z, \eta)\big|_{p_k} \quad \text{and} \quad \boldsymbol{A}_k(z_k) = \nabla_\eta K(p_k). \tag{4-49}$$

Therefore, it suffices to show the right sides in (4-49) are both zero. Recall $u(z, \eta) = \sum_{j=1}^{n} Q_{jz_j} + R[z]\eta$. We have

$$\begin{aligned}\partial_{z_{jA}} K(z, \eta)\big|_{p_k} &= \partial_{z_{jA}} E\big(u(z'(z, \eta), \eta'(z, \eta))\big)\big|_{p_k} \\ &= \operatorname{Re}\big\langle \nabla E\big(u(z'(p_k), \eta'(p_k))\big), \overline{\partial_{z_{jA}} u\big(z'(z, \eta), \eta'(z, \eta)\big)\big|_{p_k}}\big\rangle.\end{aligned}$$

By Lemma 4.8, we have

$$(z'(p_k), \eta'(p_k)) = p_k. \tag{4-50}$$

So

$$\nabla E(u(z'(p_k), \eta'(p_k))) = \nabla E(Q_{kz_k}) = 2E_{kz_k} Q_{kz_k}.$$

By Proposition 1.1 and by (4-50), for $z_k = e^{\mathrm{i}\vartheta_k}\rho_k$ we have

$$-\mathrm{i}\mathfrak{F}_* \frac{\partial}{\partial\vartheta_k}\bigg|_{p_k} = -\mathrm{i}\frac{\partial}{\partial\vartheta_k}\bigg(\sum_{j=1}^{n} Q_{jz'_j} + R[z']\eta'\bigg)\bigg|_{p_k} = -\mathrm{i}\frac{\partial}{\partial\vartheta_k} Q_{kz_k} = -\mathrm{i}\frac{\partial}{\partial\vartheta_k} e^{\mathrm{i}\vartheta_k} Q_{k\rho_k} = Q_{kz_k},$$

where the first equality follows by definition of push forward, the second by (4-50) and the third by Proposition 1.1. Similarly, by the definition of push forward, we have

$$\partial_{z_{jA}} u(z'(z,\eta),\eta'(z,\eta))\big|_{p_k} = \mathfrak{F}_* \partial_{z_{jA}}\big|_{p_k}.$$

Therefore, $b_{kjA}(z_k) = 0$ follows by

$$\partial_{z_{jA}} K(z,\eta)\big|_{p_k} = 2E_{kz_k} \operatorname{Im}\langle \mathfrak{F}_* \partial_{\vartheta_k}|_{p_k}, \overline{\mathfrak{F}_* \partial_{z_{jA}}|_{p_k}}\rangle = -E_{kz_k} \Omega_0(\partial_{\vartheta_k}, \partial_{z_{j,A}})\big|_{p_k} = 0.$$

To get $\boldsymbol{A}_k(z_k) = 0$, fix $\Xi \in \mathcal{H}_c[0]$ and set $p_{k,\Xi}(t) := (0,\dots,0,z_k,0,\dots,0;t\Xi)$. Then, for all Ξ,

$$\begin{aligned}
\operatorname{Re}\langle \nabla K(p_k), \Xi\rangle &= \frac{d}{dt} K(p_{k,\Xi}(t))\Big|_{t=0} = \frac{d}{dt} E\big(u\big(z'(p_{k,\Xi}(t)),\eta'(p_{k,\Xi}(t))\big)\big)\Big|_{t=0} \\
&= \operatorname{Re}\left\langle \nabla E(Q_{kz_k}), \overline{\frac{d}{dt} u\big(z'(p_{k,\Xi}(t)),\eta'(p_{k,\Xi}(t))\big)\Big|_{t=0}}\right\rangle \\
&= 2E_{kz_k} \operatorname{Im}\left\langle \mathfrak{F}_* \frac{\partial}{\partial \vartheta_k}\Big|_{p_k}, \overline{\mathfrak{F}_* \Xi}\right\rangle = -E_{kz_k}\Omega_0\left(\frac{\partial}{\partial\vartheta_k},\Xi\right)\Big|_{p_k} = 0 \implies \boldsymbol{A}_k(z_k) = 0. \qquad \square
\end{aligned}$$

5. Birkhoff normal form

In this section, where we search for the effective Hamiltonian, the main result is Theorem 5.9.

We consider the symplectic form Ω_0 introduced in (4-13). We introduce an index $\ell = j, \bar{j}$, for $\bar{\bar{j}} = j$ with $j = 1,\dots,n$. We write $\partial_j = \partial_{z_j}$ and $\partial_{\bar{j}} = \partial_{\bar{z}_j}$, $z_{\bar{j}} = \bar{z}_j$. With this notation, summing on j, by (4-8) and (4-34) for $\gamma_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}(|z_j|^2)$ we have

$$\Omega_0 = \mathrm{i}(1+\gamma_j(|z_j|^2))\, dz_j \wedge d\bar{z}_j + \mathrm{i}\langle d\eta, d\bar{\eta}\rangle - \mathrm{i}\langle d\bar{\eta}, d\eta\rangle. \tag{5-1}$$

Given $F \in C^1(U,\mathbb{R})$ with U an open subset of $\mathbb{C}^n \times \Sigma^c_r$, its Hamiltonian vector field X_F is defined by $i_{X_F}\Omega_0 = dF$. We have, summing on j,

$$\begin{aligned}
i_{X_F}\Omega_0 &= \mathrm{i}(1+\gamma_j(|z_j|^2))((X_F)_j\, d\bar{z}_j - (X_F)_{\bar{j}}\, dz_j) + \mathrm{i}\langle (X_F)_\eta, d\bar{\eta}\rangle - \mathrm{i}\langle (X_F)_{\bar{\eta}}, d\eta\rangle \\
&= \partial_j F\, dz_j + \partial_{\bar{j}} F\, d\bar{z}_j + \langle \nabla_\eta F, d\eta\rangle + \langle \nabla_{\bar{\eta}} F, d\bar{\eta}\rangle.
\end{aligned}$$

So, comparing the components of the two sides, we get for $1+\varpi_j(|z_j|^2) = (1+\gamma_j(|z_j|^2))^{-1}$, where $\varpi_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}(|z_j|^2)$,

$$\begin{aligned}
(X_F)_j &= -\mathrm{i}(1+\varpi_j(|z_j|^2))\partial_{\bar{j}} F, & (X_F)_\eta &= -\mathrm{i}\nabla_{\bar{\eta}} F, \\
(X_F)_{\bar{j}} &= \mathrm{i}(1+\varpi_j(|z_j|^2))\partial_j F, & (X_F)_{\bar{\eta}} &= \mathrm{i}\nabla_\eta F.
\end{aligned} \tag{5-2}$$

Given $G \in C^1(U,\mathbb{R})$ and $F \in C^1(U,\boldsymbol{E})$ with $\boldsymbol{E}$ a Banach space, we set $\{F,G\} := dF X_G$.

Definition 5.1 (normal form). Recall Definition 2.5 and, in particular, (2-13). Fix $r \in \mathbb{N}_0$. A real-valued function $Z(z,\eta)$ is in normal form if $Z = Z_0 + Z_1$, where Z_0 and Z_1 are finite sums of the following type for $\boldsymbol{l} \geq 1$:

$$Z_1(z, \boldsymbol{Z}, \eta) = \sum_{j=1}^{n} \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l} \\ \boldsymbol{m} \in \mathcal{M}_j(\boldsymbol{l})}} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}), \tag{5-3}$$

$$Z_0(z, \boldsymbol{Z}) = \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l}+1 \\ \boldsymbol{m} \in \mathcal{M}_0(\boldsymbol{l}+1)}} \boldsymbol{Z}^{\boldsymbol{m}} a_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2), \tag{5-4}$$

where $G_{j\boldsymbol{m}}(|z_j|^2) = \boldsymbol{S}^{0,0}_{r,\infty}(|z_j|^2)$, $\boldsymbol{Z}$ is as in Definition 2.2 and $a_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) = \mathcal{R}^{0,0}_{r,\infty}(|z_1|^2, \dots, |z_n|^2)$.

Remark 5.2. By Lemma 2.6, $\boldsymbol{Z}^{\boldsymbol{m}} = |z_1|^{2m_1} \cdots |z_n|^{2m_n}$ for all $\boldsymbol{m} \in \mathcal{M}_0(2N+4)$ for an $m \in \mathbb{N}_0^n$ with $2|m| = |\boldsymbol{m}|$. By Lemma 2.6 for $|\boldsymbol{m}| \leq 2N+3$, either $\sum_{a,b}(e_a - e_b)m_{ab} - e_j > 0$ or $\sum_{a,b}(e_a - e_b)m_{ab} - e_j < 0$.

For $\boldsymbol{l} \leq 2N+4$ we will consider flows associated to Hamiltonian vector fields X_χ with real-valued functions χ of the form

$$\chi = \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l}+1 \\ \boldsymbol{m} \notin \mathcal{M}_0(\boldsymbol{l}+1)}} \boldsymbol{Z}^{\boldsymbol{m}} b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) + \sum_{j=1}^{n} \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l} \\ \boldsymbol{m} \notin \mathcal{M}_j(\boldsymbol{l})}} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle B_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}) \tag{5-5}$$

with $b_{\boldsymbol{m}} = \mathcal{R}^{0,0}_{\boldsymbol{r},\infty}(|z_1|^2, \dots, |z_n|^2)$ and $B_{j\boldsymbol{m}} = \boldsymbol{S}^{0,0}_{\boldsymbol{r},\infty}(|z_j|^2)$ for some $\boldsymbol{r} \in \mathbb{N}$ defined in $B_{\mathbb{C}^n}(0, \boldsymbol{d})$ for some $\boldsymbol{d} > 0$.

The Hamiltonian vector field X_χ can be explicitly computed using (5-2). We have

$$(X_\chi)_j = (Y_\chi)_j + (\widetilde{Y}_\chi)_j, \quad (X_\chi)_\eta = -\mathrm{i} \sum_{j=1}^{n} \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l} \\ \boldsymbol{m} \notin \mathcal{M}_j(\boldsymbol{l})}} z_j \bar{\boldsymbol{Z}}^{\boldsymbol{m}} \bar{B}_{j\boldsymbol{m}}(|z_j|^2), \tag{5-6}$$

where

$$\begin{aligned}(Y_\chi)_j(z, \eta) &:= -\mathrm{i}(1 + \varpi_j(|z_j|^2)) \\ &\quad \times \Bigg[\sum_{|\boldsymbol{m}|=\boldsymbol{l}+1} b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) \partial_{\bar{j}} \boldsymbol{Z}^{\boldsymbol{m}} \\ &\qquad + \sum_{k=1}^{n} \sum_{|\boldsymbol{m}|=\boldsymbol{l}} (\langle B_{k\boldsymbol{m}}(|z_k|^2), \eta\rangle \partial_{\bar{j}}(\bar{z}_k \boldsymbol{Z}^{\boldsymbol{m}}) + \langle \bar{B}_{k\boldsymbol{m}}(|z_k|^2), \bar{\eta}\rangle \partial_{\bar{j}}(z_k \bar{\boldsymbol{Z}}^{\boldsymbol{m}})) \Bigg], \\ (\widetilde{Y}_\chi)_j(z, \eta) &:= -\mathrm{i}(1 + \varpi_j(|z_j|^2)) \Bigg[\sum_{|\boldsymbol{m}|=\boldsymbol{l}+1} \partial_{|z_j|^2} b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) z_j \boldsymbol{Z}^{\boldsymbol{m}} \\ &\qquad + \sum_{|\boldsymbol{m}|=\boldsymbol{l}} (\langle B'_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle |z_j|^2 \boldsymbol{Z}^{\boldsymbol{m}} + \langle \bar{B}'_{j\boldsymbol{m}}(|z_j|^2), \bar{\eta}\rangle z_j^2 \bar{\boldsymbol{Z}}^{\boldsymbol{m}}) \Bigg]. \end{aligned} \tag{5-7}$$

Notice that $(Y_\chi)_j = \mathcal{R}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}$, $(\widetilde{Y}_\chi)_j = \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$ and $(X_\chi)_\eta = \boldsymbol{S}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}$. We now introduce a new space.

Definition 5.3. We denote by $X_{\boldsymbol{r}}(\boldsymbol{l})$ the space formed by

$$\{(b, B) = (\{b_{\boldsymbol{m}}\}_{\boldsymbol{m}\in\mathcal{A}(\boldsymbol{l})}, \{B_{j\boldsymbol{n}}\}_{j\in 1,\dots,n,\boldsymbol{n}\in\mathcal{B}_j(\boldsymbol{l})}) : b_{\boldsymbol{m}} \in \mathbb{C},\ B_{j\boldsymbol{n}} \in \Sigma^c_{\boldsymbol{r}} \text{ and } \chi(b, B) \text{ is real valued for all } z \in B_{\mathbb{C}^n(0,\boldsymbol{d})}\},$$

where

$$\mathcal{A}(\boldsymbol{l}) := \{\boldsymbol{m} : |\boldsymbol{m}| = \boldsymbol{l}+1,\ \boldsymbol{m} \notin \mathcal{M}_0(\boldsymbol{l}+1)\},$$
$$\mathcal{B}_j(\boldsymbol{l}) := \{\boldsymbol{n} : |\boldsymbol{n}| = \boldsymbol{l},\ \boldsymbol{n} \notin \mathcal{M}_j(\boldsymbol{l}+1)\},$$

where we have assigned some order in the coordinates and where

$$\chi(b, B) = \sum_{\boldsymbol{m}\in\mathcal{A}(\boldsymbol{l})} \boldsymbol{Z}^{\boldsymbol{m}} b_{\boldsymbol{m}} + \sum_{j=1}^{n} \sum_{\boldsymbol{m}\in\mathcal{B}_j(\boldsymbol{l})} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle B_{j\boldsymbol{m}}, \eta\rangle + \text{c.c.}).$$

We give $X_{\boldsymbol{r}}(\boldsymbol{l})$ the norm

$$\|(b, B)\|_{X_{\boldsymbol{r}}(\boldsymbol{l})} = \sum_{\boldsymbol{m}\in\mathcal{A}(\boldsymbol{l})} |b_{\boldsymbol{m}}| + \sum_{j=1}^{n} \sum_{\boldsymbol{m}\in\mathcal{B}_j(\boldsymbol{l})} \|B_{j\boldsymbol{m}}\|_{\Sigma_{\boldsymbol{r}}}.$$

Set $\varrho(z) = (\varrho_1(z), \dots, \varrho_n(z))$ with $\varrho_j(z) = |z_j|^2$.

Lemma 5.4. *Consider the* χ *in* (5-5) *for fixed* $\boldsymbol{r} > 0$ *and* $\boldsymbol{l} \geq 1$, *with coefficients* $(b(\varrho(z)), B(\varrho(z)))$ *in* $C^2(B_{\mathbb{C}^n}(0, \boldsymbol{d}), X_{\boldsymbol{r}}(\boldsymbol{l}))$ *and with* $B_{j\boldsymbol{m}}(\varrho(z)) = B_{j\boldsymbol{m}}(\varrho_j(z))$. *Consider the system*

$$\dot{z}_j = (X_\chi)_j(z, \eta) \quad \text{and} \quad \dot{\eta} = (X_\chi)_\eta(z, \eta),$$

which is defined in $(t, z) \in \mathbb{R} \times B_{\mathbb{C}^n}(0, \boldsymbol{d})$ *and* $\eta \in \Sigma^c_k$ *for all* $k \in \mathbb{Z} \cap [-\boldsymbol{r}, \boldsymbol{r}]$ (*or* $\eta \in H^1 \cap \mathcal{H}_c[0]$). *Let* $\delta \in (0, \min(\boldsymbol{d}, \delta_1))$ *with* δ_1 *the constant of* Lemma 4.8. *Then the following properties hold*:

(1) *If*

$$4(\boldsymbol{l}+1)\delta \big\| \big(b(\varrho(z)), B(\varrho(z))\big) \big\|_{W^{1,\infty}(B_{\mathbb{C}^n}(0,\boldsymbol{d}), X_{\boldsymbol{r}}(\boldsymbol{l}))} < 1, \tag{5-8}$$

then, for all $k \in \mathbb{Z} \cap [-\boldsymbol{r}, \boldsymbol{r}]$, *for the flow* $\phi^t(z, \eta)$ *we have*

$$\begin{aligned} &\phi^t \in C^\infty\big((-2, 2) \times B_{\mathbb{C}^n}(0, \delta/2) \times B_{\Sigma^c_k}(0, \delta/2), B_{\mathbb{C}^n}(0, \delta) \times B_{\Sigma^c_k}(0, \delta)\big) \\ \text{and} \quad &\phi^t \in C^\infty\big((-2, 2) \times B_{\mathbb{C}^n}(0, \delta/2) \times B_{H^1\cap\mathcal{H}_c[0]}(0, \delta/2), B_{\mathbb{C}^n}(0, \delta) \times B_{H^1\cap\mathcal{H}_c[0]}(0, \delta)\big). \end{aligned} \tag{5-9}$$

In particular, for $z^t_j := z_j \circ \phi^t(z, \eta)$ *and* $\eta^t := \eta \circ \phi^t(z, \eta)$, *and in the sense of* Remark 2.10,

$$\begin{aligned} z^t_j &= z_j + S_j(t, z, \eta) \quad \text{and} \quad \eta^t = \eta + S_\eta(t, z, \eta) \\ \text{with} \quad S_j(t, z, \eta) &= \mathcal{R}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta) \quad \text{and} \quad S_\eta(t, z, \eta) = \boldsymbol{S}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta). \end{aligned} \tag{5-10}$$

(2) *We have* $S_j(t, e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} S_j(t, z, \eta)$ *and* $S_\eta(t, e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) = e^{\mathrm{i}\vartheta} S_\eta(t, z, \eta)$.

(3) *The flow* ϕ^t *is canonical, that is*, $\phi^{t*}\Omega_0 = \Omega_0$ *in* $B_{\mathbb{C}^n}(0, \delta/2) \times B_{H^1\cap\mathcal{H}_c[0]}(0, \delta/2)$.

Proof. Claim (2) is elementary. The same is true for (3), given that ϕ^t is a standard, sufficiently regular flow. In claim (1), (5-10) is a consequence of Lemma 4.9. The first part of claim (1) follows from elementary estimates such as

$$|(X_\chi)_j(z,\eta)| = |(1+\varpi_j(|z_j|^2))\partial_{\bar{j}}\chi(z,\eta)| \le (1+\|\varpi_j\|_{L^\infty(B_{\mathbb{C}}(0,\delta_0))})(\boldsymbol{l}+1)\|(b,B)\|_{W^{1,\infty}(B_{\mathbb{C}^n}(0,\delta_0),X_{\boldsymbol{r}}(\boldsymbol{l}))}\delta_0^{\boldsymbol{l}+1}$$

for $(z,\eta)\in B_{\mathbb{C}^n}(0,\delta)\times B_{\Sigma^c_{-r}}(0,\delta)$. Notice that, taking δ_0 sufficiently small in Lemma 4.6, we can arrange $\|\varpi_j\|_{L^\infty(B_{\mathbb{C}}(0,\delta_0))}<1$. We also have

$$\|(X_\chi)_\eta(z,\eta)\|_{\Sigma_{\boldsymbol{r}}} \le \|(0,B)\|_{L^\infty(B_{\mathbb{C}^n}(0,\delta_0),X_{\boldsymbol{r}}(\boldsymbol{l}))}\delta_0^{\boldsymbol{l}+1}.$$

Then if (5-8) holds we obtain (5-9). □

The main part of ϕ^t will be given by the following lemma:

Lemma 5.5. *Consider a function χ as in (5-5). For a parameter $\varrho\in[0,\infty)^n$, consider the field W_χ defined as follows (notice that $W_\chi(z,\eta,\varrho(z)) = Y_\chi(z,\eta)$):*

$$(W_\chi)_j(z,\eta,\varrho) := -\mathrm{i}(1+\varpi_j(\varrho_j)) \times\Bigg[\sum_{|\boldsymbol{m}|=\boldsymbol{l}+1} b_{\boldsymbol{m}}(\varrho)\partial_{\bar{j}}\boldsymbol{Z}^{\boldsymbol{m}} + \sum_{k=1}^{n}\sum_{|\boldsymbol{m}|=\boldsymbol{l}}\big(\langle B_{k\boldsymbol{m}}(\varrho_k),\eta\rangle\partial_{\bar{j}}(\bar{z}_k\boldsymbol{Z}^{\boldsymbol{m}}) + \langle \bar{B}_{k\boldsymbol{m}}(\varrho_k),\bar{\eta}\rangle z_k\partial_{\bar{j}}\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\big)\Bigg], \tag{5-11}$$

$$(W_\chi)_\eta(z,\eta,\varrho) := -\mathrm{i}\sum_{k=1}^{n}\sum_{|\boldsymbol{m}|=\boldsymbol{l}} z_k\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\bar{B}_{k\boldsymbol{m}}(\varrho_k).$$

Denote by $(w^t,\sigma^t)=\phi_0^t(z,\eta)$ the flow associated to the system

$$\begin{aligned}\dot{w}_j &= (W_\chi)_j(w,\sigma,\varrho(z)), & w_j(0)&=z_j,\\ \dot{\sigma} &= (W_\chi)_\sigma(w,\sigma,\varrho(z)), & \sigma(0)&=\eta.\end{aligned}\tag{5-12}$$

Let $\delta\in(0,\min(\boldsymbol{d},\delta_1))$, as in Lemma 5.4. Then the following facts hold:

(1) *If (5-8) holds, then, for $B(\varrho(z)) = (B_{j\boldsymbol{m}}(\varrho_j(z))_{j\boldsymbol{m}}$,*

$$w_j^t = z_j + T_j(t,b(\varrho(z)),B(\varrho(z)),z,\eta) \quad \textit{and} \quad \sigma^t = \eta + T_\eta(t,b(\varrho(z)),B(\varrho(z)),z,\eta), \tag{5-13}$$

$$T_j \textit{ and } T_\eta \textit{ are } C^\infty \textit{ for } (t,b,B,z,\eta)\in(-2,2)\times B_{X_{\boldsymbol{r}}}(0,c)\times B_{\mathbb{C}^n}(0,\delta)\times B_{\Sigma_{-\boldsymbol{r}}}(0,\delta) \tag{5-14}$$

with values in $\mathbb{C}$ and $\Sigma_{\boldsymbol{r}}$, respectively. Furthermore, we have

$$\begin{aligned}T_j(t,b,B,z,\eta) &= \mathcal{R}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}(t,b,B,z,\boldsymbol{Z},\eta),\\ T_\eta(t,b,B,z,\eta) &= \boldsymbol{S}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}(t,b,B,z,\boldsymbol{Z},\eta).\end{aligned}\tag{5-15}$$

(2) *We have the gauge covariance, for any fixed $\vartheta\in\mathbb{R}$,*

$$\begin{aligned}T_j(t,b,B,e^{\mathrm{i}\vartheta}z,e^{\mathrm{i}\vartheta}\eta) &= e^{\mathrm{i}\vartheta}T_j(t,b,B,z,\eta),\\ T_\eta(t,b,B,e^{\mathrm{i}\vartheta}z,e^{\mathrm{i}\vartheta}\eta) &= e^{\mathrm{i}\vartheta}T_\eta(t,b,B,z,\eta).\end{aligned}\tag{5-16}$$

(3) *Consider the Hamiltonian flow* $(z^t, \eta^t) = \phi^t(z, \eta)$ *associated to* χ; *see Lemma 5.4. Then*

$$z^t - w^t = \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta), \quad \eta^t - \sigma^t = \boldsymbol{S}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta). \tag{5-17}$$

Proof. We have (5-13)–(5-14) by standard ODE theory. For $\boldsymbol{W} = (w_i \bar{w}_j)_{i\neq j}$ like the $\boldsymbol{Z}$ in Definition 2.2,

$$\begin{aligned} w_j^t = z_j - \mathrm{i}(1+\varpi_j(\varrho_j(z)))\Bigg[&\sum_{|\boldsymbol{m}|=\boldsymbol{l}+1} b_{\boldsymbol{m}}(\varrho(z)) \int_0^t (\partial_{\bar{j}} \boldsymbol{W}^{\boldsymbol{m}})^s \, ds \\ &+ \sum_{k=1}^n \sum_{|\boldsymbol{m}|=\boldsymbol{l}} \bigg(\Big\langle B_{k\boldsymbol{m}}(\varrho_k(z)), \int_0^t \sigma^s (\partial_{\bar{j}}(\bar{w}_k \boldsymbol{W}^{\boldsymbol{m}}))^s \, ds \Big\rangle \\ &+ \Big\langle \bar{B}_{k\boldsymbol{m}}(\varrho_k(z)), \int_0^t \bar{\sigma}^s w_k^s (\partial_{\bar{j}} \overline{\boldsymbol{W}^{\boldsymbol{m}}})^s \, ds \Big\rangle \bigg) \Bigg], \end{aligned} \tag{5-18}$$

where $(\partial_{\bar{j}} \overline{\boldsymbol{W}^{\boldsymbol{m}}})^s = \partial_{\bar{j}} \overline{\boldsymbol{W}^{\boldsymbol{m}}}|_{w=w^s}$. Similarly, we have

$$\sigma^t = \eta - \mathrm{i} \sum_{k=1}^n \sum_{|\boldsymbol{m}|=\boldsymbol{l}} \bar{B}_{k\boldsymbol{m}}(\varrho_k(z)) \int_0^t w_k^s (\overline{\boldsymbol{W}^{\boldsymbol{m}}})^s \, ds. \tag{5-19}$$

Like in Lemma 4.9, we have also $\boldsymbol{W}^t = \boldsymbol{Z} + \int_0^t \mathcal{R}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}\big(s, b(\varrho(z)), B(\varrho(z)), z, \boldsymbol{Z}, \eta\big)\, ds$. We can apply Gronwall's inequality like in Lemma 4.9 in these formulas to obtain (5-15). This yields claim (1).

Next,

$$\begin{aligned} (W_\chi)_j(e^{\mathrm{i}\vartheta} w, e^{\mathrm{i}\vartheta}\sigma, \varrho(z)) &= e^{\mathrm{i}\vartheta}(W_\chi)_j(w, \sigma, \varrho(z)), \\ (W_\chi)_\eta(e^{\mathrm{i}\vartheta} w, e^{\mathrm{i}\vartheta}\sigma, \varrho(z)) &= e^{\mathrm{i}\vartheta}(W_\chi)_\eta(w, \sigma, \varrho(z)) \end{aligned}$$

yield claim (2).

Consider claim (3). Observe that (5-17) holds replacing $\boldsymbol{l}+1$ by $\boldsymbol{l}$. By (5-6), we have, for a fixed C,

$$\begin{aligned} |\dot{z} - \dot{w}| &\le |(W_\chi)_j(z, \eta) - (W_\chi)_j(w, \sigma)| + |\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta)| \\ &\le C|z-w| + C\|\eta - \sigma\|_{\Sigma_{-r}} + |\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta)|. \end{aligned}$$

Similarly, we have

$$\|\dot{\eta} - \dot{\sigma}\|_{\Sigma_r} \le \|(W_\chi)_\eta(z, \eta, \varrho(z)) - (W_\chi)_\eta(w, \sigma, \varrho(z))\|_{\Sigma_r} \le C|z-w| + C\|\eta-\sigma\|_{\Sigma_{-r}}.$$

We then conclude, by Gronwall's inequality,

$$|z^t - w^t| + \|\eta^t - \sigma^t\|_{\Sigma_r} \le |\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}(t, z, \boldsymbol{Z}, \eta)|,$$

which, along with (5-17) with $\boldsymbol{l}+1$ replaced by $\boldsymbol{l}$, yields (5-17), ending Lemma 5.5. □

Using Lemma 5.5, we expand the ϕ^1 given in Lemma 5.4.

Lemma 5.6. *Let* $(z', \eta') = \phi^1(z, \eta)$, *where* ϕ^t *is the canonical flow given in Lemma 5.4. We have*:

(1) *For* $\mathcal{T}_j(b, B, z, \eta) = \mathcal{R}^{3,2\boldsymbol{l}-1}_{\boldsymbol{r},\infty}$, $\mathcal{T}_\eta(b, B, z, \eta) = \boldsymbol{S}^{3,2\boldsymbol{l}-1}_{\boldsymbol{r},\infty}$ *and* $\mathcal{T}_j$, $\mathcal{T}_\eta$ *smooth in* (b, B, z, η),

$$\begin{aligned} z_j' &= z_j + (Y_\chi)_j(z, \eta) + \mathcal{T}_j\big(b(\varrho(z)), B(\varrho(z)), z, \eta\big) + \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}, \\ \eta' &= \eta + (X_\chi)_\eta(z, \eta) + \mathcal{T}_\eta\big(b(\varrho(z)), B(\varrho(z)), z, \eta\big) + \mathcal{S}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \end{aligned} \tag{5-20}$$

(2) *For* $\widetilde{\mathcal{T}}_j(b, B, z, \eta) = \mathcal{R}^{1,2\boldsymbol{l}}_{\boldsymbol{r},\infty}$ *smooth in* (b, B, z, η),

$$|z'_j|^2 = |z_j|^2 + \bar{z}_j(Y_\chi)_j(z, \eta) + z_j\overline{(Y_\chi)_j(z, \eta)} + \widetilde{\mathcal{T}}_j(b(\varrho(z)), B(\varrho(z)), z, \eta) + \mathcal{R}^{1,2\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \tag{5-21}$$

Remark 5.7. For $\boldsymbol{l} \geq 2$, $\mathcal{T}_j$ and $\mathcal{T}_\eta$ are absorbed in $\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$ and $\mathcal{S}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$ and do not appear in the *homological equations* in Theorem 5.9. But, if $\boldsymbol{l} = 1$, they do, although as small perturbations.

Proof. First of all, by (5-7) and by Definition 5.3, we have $\bar{z}_j(\widetilde{Y}_\chi)_j + z_j\overline{(\widetilde{Y}_\chi)_j} = 2\,\mathrm{Re}(\bar{z}_j(\widetilde{Y}_\chi)_j) = 0$. So, using the following formula to define $\mathcal{Y}_j$, we have

$$\frac{d}{dt}|z_j|^2 = \bar{z}_j(X_\chi)_j + z_j\overline{(X_\chi)_j} = \bar{z}_j(Y_\chi)_j + z_j\overline{(Y_\chi)_j} =: \mathcal{Y}_j(z, \eta). \tag{5-22}$$

Notice that $\mathcal{Y}_j$ is $\mathcal{R}^{0,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$. Therefore, we have

$$|z^s_j|^2 - |z_j|^2 = \mathcal{R}^{0,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \tag{5-23}$$

This implies

$$b(\varrho(z^s)) - b(\varrho(z)) = \mathcal{R}^{0,\boldsymbol{l}+1}_{\boldsymbol{r},\infty} \quad \text{and} \quad B(\varrho(z^s)) - B(\varrho(z)) = \mathcal{S}^{0,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \tag{5-24}$$

Similarly — see right before (5-2) — we have

$$\varpi_j(|z^s_j|^2) - \varpi_j(|z_j|^2) = \mathcal{R}^{2,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \tag{5-25}$$

Now we show (1). By (5-6) and (5-11), using (5-24) and (5-25), we have

$$(Y_\chi)_j(z^s, \eta^s) - (W_\chi)_j(z^s, \eta^s, \varrho(z)) = \mathcal{R}^{1,2\boldsymbol{l}+1}_{\boldsymbol{r},\infty}. \tag{5-26}$$

By (5-6), (5-10), (5-17) and (5-26), we have

$$\begin{aligned}
z'_j &= z_j + \int_0^1 (W_\chi)_j(z^s, \eta^s, \varrho(z))\, ds + \int_0^1 \big((Y_\chi)_j(z^s, \eta^s) - (W_\chi)_j(z^s, \eta^s, \varrho(z))\big)\, ds + \int_0^1 (\widetilde{Y}_\chi)_j(z^s, \eta^s)\, ds \\
&= z_j + \int_0^1 (W_\chi)_j(w^s + \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}, \sigma^s + S^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}, \varrho(z))\, ds + \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty} \\
&= z_j + \int_0^1 (W_\chi)_j(w^s, \sigma^s, \varrho(z))\, ds + \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty} \\
&= z_j + (W_\chi)_j(z, \eta, \varrho(z)) + \mathcal{T}_j + \mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty},
\end{aligned}$$

where $\mathcal{T}_j = \int_0^1 (W_\chi)_j(w^s, \sigma^s, \varrho(z))\, ds - (W_\chi)_j(z, \eta, \varrho(z))$ and the last $\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$ in the second line is different from the $\mathcal{R}^{1,\boldsymbol{l}+1}_{\boldsymbol{r},\infty}$ in the third line. Finally, by Lemma 5.5(1) and the fact $(W_\chi)_j = \mathcal{R}^{1,\boldsymbol{l}}_{\boldsymbol{r},\infty}$, we have $\mathcal{T}_j = \mathcal{R}^{1,2\boldsymbol{l}-1}_{\boldsymbol{r},\infty}$ with $\mathcal{T}_j$ smooth in (t, b, B, z, η). The argument for η' is similar.

We next show (2). Set $\widetilde{\mathcal{Y}}_j(z, \eta, \varrho) := \bar{z}_j(W_\chi)_j(z, \eta, \varrho) + z_j\overline{(W_\chi)_j(z, \eta, \varrho)}$. As in (5-23)–(5-24), we have

$$\widetilde{\mathcal{Y}}_j(z^s, \eta^s, \varrho(z)) - \mathcal{Y}_j(z^s, \eta^s) = \mathcal{R}^{0,2\boldsymbol{l}+2}_{\boldsymbol{r},\infty},$$

where $\mathcal{Y}_j$ is as defined in (5-22). So we have

$$|z_j'|^2 = |z_j|^2 + \int_0^1 \mathcal{Y}_j(z^s, \eta^s)\, ds = |z_j|^2 + \int_0^1 \widetilde{\mathcal{Y}}_j(z^s, \eta^s, \varrho(z))\, ds + \mathcal{R}^{0,2\boldsymbol{l}+2}_{\boldsymbol{r},\infty}$$
$$= |z_j|^2 + \int_0^1 \widetilde{\mathcal{Y}}_j(w^s, \sigma^s, \varrho(z))\, ds + \mathcal{R}^{1,2\boldsymbol{l}+1}_{\boldsymbol{r},\infty} = |z_j|^2 + \widetilde{\mathcal{Y}}_j(z,\eta) + \widetilde{\mathcal{T}}_j + \mathcal{R}^{1,2\boldsymbol{l}+1}_{\boldsymbol{r},\infty},$$

where $\widetilde{\mathcal{T}}_j = \int_0^1 \widetilde{\mathcal{Y}}_j(w^s, \sigma^s, \varrho(z))\, ds - \widetilde{\mathcal{Y}}_j(z,\eta)$. As in (1), $\widetilde{\mathcal{T}}_j = \mathcal{R}^{1,2\boldsymbol{l}}_{\boldsymbol{r},\infty}$ and $\widetilde{T}$ is C^∞ for (b, B, z, η). $\square$

After a coordinate change $\phi = \phi^1$ as in Lemma 5.4 the Hamiltonian expands like in (4-45).

Lemma 5.8 (structure lemma). *Consider a function K which admits an expansion as in (4-45), defined for $(z,\eta) \in B_{\mathbb{C}^n}(0,\delta) \times (B_{H^1}(0,\delta) \cap \mathcal{H}_c[0])$ for some small $\delta > 0$ and with r_1 replaced by an r'. Suppose also that the $l = 0$ terms in the third line are zero. Consider a function χ such as in (5-5) with $1 \le \boldsymbol{l} \le 2N+4$, with $\|(b, B)\|_{W^{1,\infty}(B_{\mathbb{C}^n}(0,\delta), X_r(\boldsymbol{l}))} \le \underline{C}$ and with $\underline{C}$ a preassigned number. Suppose also that $2c_2(2N+4)\delta\underline{C} < 1$ with c_2 the constant of Lemma 5.4. Denote by $\phi = \phi^1$ the corresponding flow. Then Lemma 5.4(1)–(3) hold and, for $(z,\eta) \in B_{\mathbb{C}^n}(0,\delta/2) \times (B_{H^1}(0,\delta/2) \cap \mathcal{H}_c[0])$, $r = r' - 2$ and $\boldsymbol{Z}$ as in Definition 2.2, we have an expansion*

$$\begin{aligned}
K \circ \phi(z,\eta) &= H_2(z,\eta) + \sum_{j=1}^n \lambda_j(|z_j|^2) + \sum_{l=1}^{2N+3} \sum_{|\boldsymbol{m}|=l+1} \boldsymbol{Z}^{\boldsymbol{m}} a_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) \\
&\quad + \sum_{j=1}^n \sum_{l=1}^{2N+3} \sum_{|\boldsymbol{m}|=l} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}) + \mathcal{R}^{1,2}_{r,\infty}(z,\eta) + \mathcal{R}^{0,2N+5}_{r,\infty}(z, \boldsymbol{Z}, \eta) \\
&\quad + \operatorname{Re}\langle \boldsymbol{S}^{0,2N+4}_{r,\infty}(z, \boldsymbol{Z}, \eta), \bar{\eta}\rangle + \sum_{i+j=2} \sum_{|\boldsymbol{m}|\le 1} \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{2\boldsymbol{m}ij}(z,\eta), \eta^i \bar{\eta}^j\rangle \\
&\quad + \sum_{d+c=3} \sum_{i+j=d} \langle G_{dij}(z,\eta), \eta^i \bar{\eta}^j\rangle \mathcal{R}^{0,c}_{r,\infty}(z,\eta) + E_P(\eta),
\end{aligned} \tag{5-27}$$

where $G_{j\boldsymbol{m}}$, $G_{2\boldsymbol{m}ij}$ and G_{dij} are $\boldsymbol{S}^{0,0}_{r,\infty}$ and the $a_{\boldsymbol{m}}$ are $\mathcal{R}^{0,0}_{\infty,\infty}$. For $|\boldsymbol{m}| = 0$, we have $G_{2\boldsymbol{m}ij}(z,\eta) = G_{2\boldsymbol{m}ij}(z)$ are the functions in (3-4) and the $\lambda_j(|z_j|^2)$ are the same as those of (4-45). Furthermore, the term $\mathcal{R}^{1,2}_{r,\infty}(z,\eta)$ in (5-27) satisfies $\mathcal{R}^{1,2}_{r,\infty}(e^{\mathrm{i}\vartheta} z, e^{\mathrm{i}\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r,\infty}(z,\eta)$.

Proof. Like in Lemma 4.10, we consider the expansion (4-45) for $K(z', \eta')$, and substitute the formulas $z_j' = z_j + S_j(z,\eta)$ and $\eta' = \eta + S_\eta(z,\eta)$. Proceeding like in Lemma 4.10, we have

$$\mathcal{R}^{1,2}_{r',\infty}(z', \eta') = \mathcal{R}^{1,2}_{r',\infty}(z,\eta) + \mathcal{R}^{1,2N+5}_{r',\infty}(z, \boldsymbol{Z}, \eta) + \operatorname{Re}\langle \boldsymbol{S}^{1,2N+4}_{r',\infty}(z, \boldsymbol{Z}, \eta), \bar{\eta}\rangle + \mathfrak{S}, \tag{5-28}$$

where $\mathfrak{S}$ consists of terms like in the second and third sums of (5-27).

Similarly, for a $\widetilde{\mathfrak{S}}$ like $\mathfrak{S}$, we have

$$\begin{aligned}
\langle H\eta', \bar{\eta}'\rangle &= \langle H\eta, \bar{\eta}\rangle + \mathcal{R}^{1,\boldsymbol{l}+1}_{r'-2,\infty}(z, \boldsymbol{Z}, \eta) \\
&= \langle H\eta, \bar{\eta}\rangle + \mathcal{R}^{1,\boldsymbol{l}+1}_{r'-2,\infty}(z,\eta) + \mathcal{R}^{1,\boldsymbol{l}+1}_{r'-2,\infty}(z, \boldsymbol{Z}) + \operatorname{Re}\langle \boldsymbol{S}^{1,\boldsymbol{l}}_{r'-2,\infty}(z, \boldsymbol{Z}, \eta), \bar{\eta}\rangle \\
&= \langle H\eta, \bar{\eta}\rangle + \mathcal{R}^{1,\boldsymbol{l}+1}_{r'-2,\infty}(z,\eta) + \mathcal{R}^{1,2N+5}_{r'-2,\infty}(z, \boldsymbol{Z}, \eta) + \operatorname{Re}\langle \boldsymbol{S}^{1,2N+4}_{r'-2,\infty}(z, \boldsymbol{Z}, \eta), \bar{\eta}\rangle + \widetilde{\mathfrak{S}}.
\end{aligned} \tag{5-29}$$

Consider a $\lambda_j(|z_j|^2)$ in (4-45). Then, by (5-21), we have

$$\lambda(|z_j'|^2) = \lambda(|z_j|^2 + \mathcal{R}^{0,\boldsymbol{l}+1}_{r,\infty}(z,\boldsymbol{Z},\eta)) = \mu(|z_j|^2) + \mathcal{R}^{1,\boldsymbol{l}+1}_{r,\infty}(z,\boldsymbol{Z},\eta). \tag{5-30}$$

The latter admits an expansion like in (4-46) and what follows it.

The term $\mathcal{R}^{1,2}_{r,\infty}(z,\eta)$ in the second line of (5-27) is either the first in the right-hand side in (5-28) for $l > 1$ in Lemma 4.8, or the sum of that with the $\mathcal{R}^{1,l+1}_{r'-2,\infty}(z,\eta)$ originating from (5-29)–(5-30) for $l = 1$ in Lemma 4.8. In either case it satisfies $\mathcal{R}^{1,2}_{r,\infty}(e^{i\vartheta}z, e^{i\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r,\infty}(z,\eta)$. Other terms in (4-45) computed at (z',η') by similar elementary expansions are similarly absorbed in (5-27). □

All of the above lemmas are preparation for the following result, which will give us an effective Hamiltonian by picking $\iota = 2N+4$.

Theorem 5.9 (Birkhoff normal form). *For any $\iota \in \mathbb{N} \cap [2, 2N+4]$ there is a $\delta_\iota > 0$, a polynomial χ_ι as in (5-5) with $\boldsymbol{l} = \iota$, $\boldsymbol{d} = \delta_\iota$ and $\boldsymbol{r} = r_\iota = r_0 - 2(\iota+1)$ such that, for all $k \in \mathbb{Z} \cap [-r(\iota), r(\iota)]$, we have for each χ_ι a flow (for δ_1 the constant in Lemma 4.10)*

$$\begin{aligned} &\phi^t_\iota \in C^\infty\big((-2,2) \times B_{\mathbb{C}^n}(0,\delta_\iota) \times B_{\Sigma^c_k}(0,\delta_\iota), B_{\mathbb{C}^n}(0,\delta_{\iota-1}) \times B_{\Sigma^c_k}(0,\delta_{\iota-1})\big) \\ \text{and}\quad &\phi^t_\iota \in C^\infty\big((-2,2) \times B_{\mathbb{C}^n}(0,\delta_\iota) \times B_{H^1\cap\mathcal{H}_c[0]}(0,\delta_\iota), B_{\mathbb{C}^n}(0,\delta_{\iota-1}) \times B_{H^1\cap\mathcal{H}_c[0]}(0,\delta_{\iota-1})\big) \end{aligned} \tag{5-31}$$

and such that, for $\mathfrak{F}^{(\iota)} := \mathfrak{F} \circ \phi_2 \circ \cdots \circ \phi_\iota$ with $\mathfrak{F}$ the transformation in Lemma 4.8 and $\phi_j = \phi^1_\iota$, for $(z,\eta) \in B_{\mathbb{C}^n}(0,\delta_\iota) \times (B_{H^1}(0,\delta_\iota) \cap \mathcal{H}_c[0])$ and for $\boldsymbol{Z}$ as in Definition 2.2, we have

$$\begin{aligned} H^{(\iota)}(z,\eta) &:= E \circ \mathfrak{F}^{(\iota)}(z,\eta) \\ &= H_2(z,\eta) + \sum_{j=1}^n \lambda_j(|z_j|^2) + Z^{(\iota)}(z,\boldsymbol{Z},\eta) + \sum_{l=\iota}^{2N+3} \sum_{|\boldsymbol{m}|=l+1} \boldsymbol{Z}^{\boldsymbol{m}} a^{(\iota)}_{\boldsymbol{m}}(|z_1|^2,\dots,|z_n|^2) \\ &\quad + \sum_{j=1}^n \sum_{l=\iota}^{2N+3} \sum_{|\boldsymbol{m}|=l} (\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G^{(\iota)}_{j\boldsymbol{m}}(|z_j|^2),\eta\rangle + \text{c.c.}) + \mathcal{R}^{1,2}_{r_\iota,\infty}(z,\eta) + \mathcal{R}^{0,2N+5}_{r_\iota,\infty}(z,\boldsymbol{Z},\eta) \\ &\quad + \operatorname{Re}\langle \boldsymbol{S}^{0,2N+4}_{r_\iota,\infty}(z,\boldsymbol{Z},\eta),\bar{\eta}\rangle + \sum_{i+j=2} \sum_{|\boldsymbol{m}|\le 1} \boldsymbol{Z}^{\boldsymbol{m}} \langle G^{(\iota)}_{2\boldsymbol{m}ij}(z,\eta),\eta^i\bar{\eta}^j\rangle \\ &\quad + \sum_{d+c=3} \sum_{i+j=d} \langle G^{(\iota)}_{dij}(z,\eta),\eta^i\bar{\eta}^j\rangle \mathcal{R}^{0,c}_{r_\iota,\infty}(z,\eta) + E_P(\eta), \end{aligned} \tag{5-32}$$

where, for coefficients like in Definition 5.1 for $(r,m) = (r_\iota,\infty)$,

$$Z^{(\iota)} = \sum_{\boldsymbol{m}\in\mathcal{M}_0(\iota)} \boldsymbol{Z}^{\boldsymbol{m}} a_{\boldsymbol{m}}(|z_1|^2,\dots,|z_n|^2) + \sum_{j=1}^n \bigg(\sum_{\boldsymbol{m}\in\mathcal{M}_j(\iota-1)} \bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{j\boldsymbol{m}}(|z_j|^2),\eta\rangle + \text{c.c.} \bigg). \tag{5-33}$$

We have $\mathcal{R}^{1,2}_{r_\iota,\infty} = \mathcal{R}^{1,2}_{r_2,\infty}$ and $\mathcal{R}^{1,2}_{r_2,\infty}(e^{i\vartheta}z, e^{i\vartheta}\eta) \equiv \mathcal{R}^{1,2}_{r_2,\infty}(z,\eta)$.

In particular, we have, for $\delta_f := \delta_{2N+4}$ and for the δ_0 in Lemma 4.6,

$$\mathfrak{F}^{(2N+4)}\big(B_{\mathbb{C}^n}(0,\delta_f) \times (B_{H^1}(0,\delta_f) \cap \mathcal{H}_c[0])\big) \subset B_{\mathbb{C}^n}(0,\delta_0) \times (B_{H^1}(0,\delta_0) \cap \mathcal{H}_c[0]) \tag{5-34}$$

with $\mathfrak{F}|_{B_{\mathbb{C}^n}(0,\delta_f)\times(B_{H^1}(0,\delta_f)\cap\mathcal{H}_c[0])}$ a diffeomorphism between its domain and an open neighborhood of the origin in $\mathbb{C}^n\times(H^1\cap\mathcal{H}_c[0])$.

Furthermore, for $r=r_0-4N-10$, there is a pair $\mathcal{R}^{1,1}_{r,\infty}$ and $\boldsymbol{S}^{1,1}_{r,\infty}$ such that, for $(z',\eta')=\mathfrak{F}^{(2N+4)}(z,\eta)$,

$$z'=z+\mathcal{R}^{1,1}_{r,\infty}(z,\boldsymbol{Z},\eta)\quad\text{and}\quad \eta'=\eta+\boldsymbol{S}^{1,1}_{r,\infty}(z,\boldsymbol{Z},\eta). \tag{5-35}$$

By taking all the $\delta_\iota>0$ sufficiently small, we can assume that all the symbols in the proof, i.e., the symbols in (5-35) and the symbols in the expansions (5-32), satisfy the estimates of Definitions 2.8 and 2.9 for $|z|<\delta_\iota$ and $\|\eta\|_{\Sigma_{r(\iota)}}<\delta_\iota$ for their respective ι.

Proof. Notice that the functional K in Lemma 4.10 satisfies the case $\iota=1$. The proof will be by induction on ι. We assume that $H^{(\iota)}$ satisfies the statement for some $\iota\geq 1$ and prove that there is a $\phi_{\iota+1}$ such that $H^{(\iota+1)}:=H^{(\iota)}\circ\phi_{\iota+1}$ satisfies the statement for $\iota+1$. We consider the representation (5-27) for $H^{(\iota)}$, which is guaranteed by Lemma 5.8. Using (5-27), we set $\boldsymbol{h}=H^{(\iota)}(z,\boldsymbol{Z},\eta)$, interpreting $(z,\boldsymbol{Z},\eta)$ as independent variables. Then we have, for $\boldsymbol{l}=\iota$,

$$a^{(\boldsymbol{l})}_{\boldsymbol{m}}(|z_1|^2,\dots,|z_n|^2)=\frac{1}{\boldsymbol{m}!}\partial^{\boldsymbol{m}}_{\boldsymbol{Z}}\boldsymbol{h}\Big|_{(z,\eta,\boldsymbol{Z})=(z;0,0)},\qquad |\boldsymbol{m}|\leq 2N+4, \tag{5-36}$$

$$\bar z_j G^{(\boldsymbol{l})}_{j\boldsymbol{m}}(|z_j|^2)=\frac{1}{\boldsymbol{m}!}\partial^{\boldsymbol{m}}_{\boldsymbol{Z}}\nabla_\eta\boldsymbol{h}\Big|_{(z,\eta,\boldsymbol{Z})=(0,\dots,z_j,0,\dots,0;0,0)},\qquad |\boldsymbol{m}|\leq 2N+3. \tag{5-37}$$

The inductive hypothesis on $H^{(\iota)}$ is a statement on the Taylor coefficients in (5-36)–(5-37), that is, that, for $\boldsymbol{l}=\iota$ (see Definition 2.5 and Remark 5.2),

$$\partial^{\boldsymbol{m}}_{\boldsymbol{Z}}\boldsymbol{h}\big|_{(z,\eta,\boldsymbol{Z})=(z;0,0)}=0\quad\text{for all }\boldsymbol{m}\notin\mathcal{M}_0(\boldsymbol{l}), \tag{5-38}$$

$$\partial^{\boldsymbol{m}}_{\boldsymbol{Z}}\nabla_\eta\boldsymbol{h}\big|_{(z,\eta,\boldsymbol{Z})=(0,\dots,z_j,0,\dots,0;0,0)}=0\quad\text{for all }(j,\boldsymbol{m})\text{ with }\boldsymbol{m}\notin\mathcal{M}_j(\boldsymbol{l}-1). \tag{5-39}$$

We consider now an as yet unknown χ as in (5-5) with $\boldsymbol{l}=\iota$, $\boldsymbol{r}=r_\iota$ and a yet to be determined $\boldsymbol{d}=\delta>0$. Set $\phi:=\phi^1$, where ϕ^t is the flow of Lemma 5.4. We are seeking χ such that $H^{(\iota)}\circ\phi$ satisfies the conclusions of Theorem 5.9 for $\iota+1$, i.e., that using Lemma 5.8 again and setting this time $\boldsymbol{h}=(H^{(\iota)}\circ\phi)(z,\eta,\boldsymbol{Z})$, we will have (5-38)–(5-39) for $\boldsymbol{l}=\iota+1$. Notice that, for any χ, (5-38)–(5-39) are automatically true for $\boldsymbol{l}=\iota$. This is because $H^{(\iota)}(z,\eta,\boldsymbol{Z})$ and $(H^{(\iota)}\circ\phi)(z,\eta,\boldsymbol{Z})$ have the same derivatives in (5-36) for $|\boldsymbol{m}|\leq\iota$, and in (5-37) for $|\boldsymbol{m}|\leq\iota-1$. So it is enough to consider (5-38) for $|\boldsymbol{m}|=\iota+1$ and (5-39) for $|\boldsymbol{m}|=\iota$. This will be true for a specific choice of χ whose coefficients solve the *homological equations*, which we set up in the sequel.

By (5-20) and $G^{(\iota)}_{\boldsymbol{20}ij}(z,\eta)=G_{\boldsymbol{20}ij}(z)$, we have

$$\begin{aligned}H^{(\iota)}(z',\eta')=H_2(z',\eta')+\sum_{j=1}^n\lambda_j(|z'_j|^2)+Z^{(\iota)}(z',\boldsymbol{Z}',\eta')+\mathcal{R}^{1,2}_{r,\infty}(z',\eta')+\sum_{i+j=2}\langle G_{\boldsymbol{20}ij}(z'),\eta'^i\bar\eta'^j\rangle\\ +(*)+\sum_{|\boldsymbol{m}|=\iota+1}\boldsymbol{Z}^{\boldsymbol{m}}a^{(\iota)}_{\boldsymbol{m}}(|z|^2)+\sum_{j=1}^n\sum_{|\boldsymbol{m}|=\iota}(\bar z_j\boldsymbol{Z}^{\boldsymbol{m}}\langle G^{(\iota)}_{j\boldsymbol{m}}(|z_j|^2),\eta\rangle+\text{c.c.}),\end{aligned} \tag{5-40}$$

where $\boldsymbol{h} := (*)(z, \eta, \boldsymbol{Z})$ satisfies (5-38)–(5-39) for $\boldsymbol{l} = \iota + 1$. In the sequel, we will use $(*)$ with this meaning. Let $(z', \eta') = \phi(z, \eta)$. We have

$$\begin{aligned}\sum_{j=1}^{n} e_j(\bar z_j (Y_\chi)_j(z,\eta) + z_j (Y_\chi)_{\bar j}(z,\eta)) \\ &= \sum_{|\boldsymbol{m}|=\iota+1} \mathrm{i}\tilde{\boldsymbol{e}}\cdot(\mu(\boldsymbol{m}) - \nu(\boldsymbol{m})) b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) \boldsymbol{Z}^{\boldsymbol{m}} \\ &\quad + \sum_j \sum_{|\boldsymbol{m}|=\iota} \big(\mathrm{i}\tilde{\boldsymbol{e}}\cdot(\tilde\mu_j(\boldsymbol{m}) - \tilde\nu_j(\boldsymbol{m}))\langle B_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle \bar z_j \boldsymbol{Z}^{\boldsymbol{m}} + \text{c.c.}\big) \end{aligned} \tag{5-41}$$

for

$$\begin{aligned} \boldsymbol{Z}^{\boldsymbol{m}} &= z^{\mu(\boldsymbol{m})} \bar z^{\nu(\boldsymbol{m})}, \\ \bar z_j \boldsymbol{Z}^{\boldsymbol{m}} &= z^{\tilde\mu_j(\boldsymbol{m})} \bar z^{\tilde\nu_j(\boldsymbol{m})}, \\ \tilde{\boldsymbol{e}}(z) &:= \big(e_1(1+\varpi_1(|z_1|^2)), \dots, e_n(1+\varpi_n(|z_n|^2))\big), \end{aligned} \tag{5-42}$$

and, summing on repeated indexes,

$$\langle H\eta, (X_\chi)_{\bar\eta}(z,\eta)\rangle + \langle H(X_\chi)_\eta(z,\eta), \bar\eta\rangle = \mathrm{i}\bar z_j \boldsymbol{Z}^{\boldsymbol{m}} \langle H B_{j,\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.} \tag{5-43}$$

So, by Lemma 5.6, (5-41)–(5-43) and using the notation in (5-42), we have

$$\begin{aligned} H_2(z', \eta') &= \sum_{j=1}^{n} e_j |z'_j|^2 + \langle H\eta', \bar\eta'\rangle \\ &= H_2(z,\eta) + \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l}+1 \\ \boldsymbol{m}\notin \mathcal{M}_0(\boldsymbol{l}+1)}} \mathrm{i}\tilde{\boldsymbol{e}}\cdot(\mu(\boldsymbol{m}) - \nu(\boldsymbol{m})) b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) \boldsymbol{Z}^{\boldsymbol{m}} \\ &\quad + \sum_j \sum_{\substack{|\boldsymbol{m}|=\boldsymbol{l} \\ \boldsymbol{m}\notin \mathcal{M}_j(\boldsymbol{l})}} \big(\mathrm{i}\big\langle\big(\tilde{\boldsymbol{e}}\cdot(\tilde\mu_j(\boldsymbol{m}) - \tilde\nu_j(\boldsymbol{m})) + H\big) B_{j\boldsymbol{m}}(|z_j|^2), \eta\big\rangle \bar z_j \boldsymbol{Z}^{\boldsymbol{m}} + \text{c.c.}\big) \\ &\quad + \mathcal{R}^{2,2\iota}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) + (*), \end{aligned} \tag{5-44}$$

where c.c. refers only to the third line and, in the last line,

$$\mathcal{R}^{2,2\iota}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) = \sum_{j=1}^{n} e_j \widetilde{\mathcal{T}}_j + \langle H\eta, \overline{\mathcal{T}}_\eta\rangle + \langle H\mathcal{T}_\eta, \bar\eta\rangle + \langle H\mathcal{T}_\eta, \overline{\mathcal{T}}_\eta\rangle,$$

where here and in the sequel of this proof we abuse notation, denoting by (b, B) the element in $X_r(\iota)$ — see Definition 5.3 — with entries $b_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2)$ and $B_{j\boldsymbol{m}}(|z_j|^2)$. The term $\mathcal{R}^{2,2\iota}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta)$ can be absorbed in $(*)$ if $\iota \ge 2$, but if $\iota = 1$ needs to be considered explicitly. By $\lambda_j(|z_j|^2) = \mathcal{R}^{2,0}_{\infty,\infty}$ and (5-21), we have

$$\lambda_j(|z'_j|^2) = \lambda_j(|z_j|^2) + \mathcal{R}^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) + (*). \tag{5-45}$$

Next, we claim

$$Z^{(\iota)}(z', \boldsymbol{Z}', \eta') = Z^{(\iota)}(z, \boldsymbol{Z}, \eta) + \mathcal{R}^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) + (*). \tag{5-46}$$

Let us take a term $\boldsymbol{Z}^{\boldsymbol{m}} a_{\boldsymbol{m}}(\varrho(z))$ in the first sum in (5-33). Notice that, by Lemma 2.6, we have necessarily $|\boldsymbol{m}| \geq 2$. Furthermore, by (5-21) it is easy to see that we can omit the factor $a_{\boldsymbol{m}}(\varrho(z))$. For definiteness, let $\boldsymbol{Z}^{\boldsymbol{m}} = |z_1|^2 |z_2|^2$ (so $|\boldsymbol{m}| = 2$; the case $|\boldsymbol{m}| > 2$ is simpler). By (5-21) we have

$$|z_1'|^2 |z_2'|^2 = (|z_1|^2 + \mathcal{R}^{0,\iota+1}_{\boldsymbol{r},\infty})(|z_2|^2 + \mathcal{R}^{0,\iota+1}_{\boldsymbol{r},\infty}) = |z_1|^2|z_2|^2 + R^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta),$$

where we used information, such as $\widetilde{\mathcal{T}}_j = \mathcal{R}^{1,2\iota}_{\boldsymbol{r},\infty}$, contained in Lemma 5.6 and the fact, easy to check, that $\bar{z}_j (Y_\chi)_j(z,\eta) + z_j (Y_\chi)_{\bar{j}}(z,\eta) = R^{0,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta)$.

To complete the proof of (5-46) let us take now a term of the form $\bar{z}_2 \boldsymbol{Z}^{\boldsymbol{m}} \langle G(|z_2|^2), \eta\rangle$. Here we can write $G = G(|z_2|^2)$, ignoring the dependence on $|z_2|^2$ and we can focus on $|\boldsymbol{m}| = 1$. For definiteness, let $\boldsymbol{Z}^{\boldsymbol{m}} = z_1 \bar{z}_2$. By Lemma 5.6,

$$z_1'(\bar{z}_2')^2 \langle G, \eta'\rangle = (z_1 + \mathcal{R}^{1,\iota}_{\boldsymbol{r},\infty})(\bar{z}_2 + \mathcal{R}^{1,\iota}_{\boldsymbol{r},\infty})^2 \langle G, \eta + S^{1,\iota}_{\boldsymbol{r},\infty}\rangle,$$

which for $\iota > 1$ is of the form $z_1 \bar{z}_2^2 \langle G, \eta\rangle + (*)$, and for $\iota = 1$, using (5-20), yields (5-46).

By Lemma 5.4(1) and $d_\eta \mathcal{R}^{1,2}_{r,\infty}(z,\eta) \cdot \boldsymbol{S}^{1,\iota}_{\boldsymbol{r},\infty}(b, B, z, \eta) = \mathcal{R}^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta)$, we get

$$\begin{aligned}
\mathcal{R}^{1,2}_{r,\infty}(z', \eta') &= \mathcal{R}^{1,2}_{r,\infty}(z, \eta') + (*) \\
&= \mathcal{R}^{1,2}_{r,\infty}(z, \eta) + (*) + \int_0^1 d_\eta \mathcal{R}^{1,2}_{r,\infty}(z, \eta + \tau \boldsymbol{S}^{1,\iota}_{\boldsymbol{r},\infty}(b, B, z, \eta)) \cdot \boldsymbol{S}^{1,\iota}_{\boldsymbol{r},\infty}(b, B, z, \eta)\, d\tau \\
&= \mathcal{R}^{1,2}_{r,\infty}(z, \eta) + d_\eta \mathcal{R}^{1,2}_{r,\infty}(z, \eta) \cdot \boldsymbol{S}^{1,\iota}_{\boldsymbol{r},\infty}(b, B, z, \eta) + (*).
\end{aligned} \tag{5-47}$$

Like in (5-47) and using (5-20) and $G_{\boldsymbol{20}ij}(z) = \mathcal{R}^{2,0}_{\infty,\infty}(z)$ — see (3-4) — we have

$$\begin{aligned}
\sum_{i+j=2} \langle G_{\boldsymbol{20}ij}(z'), \eta'^i \bar{\eta}'^j \rangle &= \sum_{i+j=2} \langle G_{\boldsymbol{20}ij}(z), \eta'^i \bar{\eta}'^j \rangle + (*) \\
&= \sum_{i+j=2} \langle G_{\boldsymbol{20}ij}(z), \eta^i \bar{\eta}^j \rangle + \mathcal{R}^{3,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) + (*).
\end{aligned} \tag{5-48}$$

Therefore, we seek χ_ι such that the following holds, with $\varrho(z) = (|z_1|^2, \dots, |z_n|^2)$ and the notation in (5-42):

$$\begin{aligned}
(*) = &\sum_{\substack{|\boldsymbol{m}|=\iota+1 \\ \boldsymbol{m} \notin \mathcal{M}_0(\iota+1)}} \mathrm{i}\tilde{\boldsymbol{e}} \cdot (\mu(\boldsymbol{m}) - \nu(\boldsymbol{m})) b_{\boldsymbol{m}}(\varrho(z)) \boldsymbol{Z}^{\boldsymbol{m}} \\
&+ \sum_j \sum_{\substack{|\boldsymbol{m}|=\iota \\ \boldsymbol{m} \notin \mathcal{M}_j(\iota)}} \big(\mathrm{i}\langle (\tilde{\boldsymbol{e}} \cdot (\mu_j(\boldsymbol{m}) - \nu_j(\boldsymbol{m})) + H) B_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle \bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} + \text{c.c.}\big) + \mathcal{R}^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta) \\
&+ \sum_{\substack{|\boldsymbol{m}|=\iota+1 \\ \boldsymbol{m} \notin \mathcal{M}_0(\iota+1)}} \boldsymbol{Z}^{\boldsymbol{m}} a^{(\iota)}_{\boldsymbol{m}}(\varrho(z)) + \sum_{j=1}^n \sum_{\substack{|\boldsymbol{m}|=\iota \\ \boldsymbol{m} \notin \mathcal{M}_j(\iota)}} \big(\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G^{(\iota)}_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}\big).
\end{aligned} \tag{5-49}$$

By a Taylor expansion, we can write

$$
\begin{aligned}
&\mathscr{R}^{2,\iota+1}_{\boldsymbol{r},\infty}(b, B, z, \boldsymbol{Z}, \eta)\\
&\quad = (*) + \sum_{\substack{|\boldsymbol{m}|=\iota+1\\ \boldsymbol{m}\notin\mathcal{M}_0(\iota+1)}} \boldsymbol{Z}^{\boldsymbol{m}} \alpha_{\boldsymbol{m}}(b, B, \varrho(z))\\
&\qquad + \sum_{j=1}^{n} \sum_{\substack{|\boldsymbol{m}|=\iota\\ \boldsymbol{m}\notin\mathcal{M}_j(\iota)}} \big(\bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle \Gamma_{j\boldsymbol{m}}(b(0,\dots,|z_j|^2,0,\dots,0), B(0,\dots,|z_j|^2,0,\dots,0), |z_j|^2), \eta\rangle + \text{c.c.}\big),
\end{aligned}
$$

where $\alpha_{\boldsymbol{m}}(b, B, \varrho(z)) = \mathscr{R}^{1,0}_{\boldsymbol{r},\infty}(b, B, \varrho(z))$ and

$$
\begin{aligned}
&\Gamma_{j\boldsymbol{m}}(b(0,\dots,|z_j|^2,0,\dots,0), B(0,\dots,|z_j|^2,0,\dots,0), |z_j|^2)\\
&\qquad\qquad = \boldsymbol{S}^{1,0}_{\boldsymbol{r},\infty}(b(0,\dots,|z_j|^2,0,\dots,0), B(0,\dots,|z_j|^2,0,\dots,0), |z_j|^2).
\end{aligned}
$$

Furthermore, by (5-42) and $\varpi_j(|z_j|^2) = \mathscr{R}^{2,0}_{r_0,\infty}(|z_j|^2)$, the sum in the second line of (5-49) has an expansion

$$
\sum_{j} \sum_{\substack{|\boldsymbol{m}|=\iota\\ \boldsymbol{m}\notin\mathcal{M}_j(\iota)}} \big(\mathrm{i}\langle(\boldsymbol{e}\cdot(\mu_j(\boldsymbol{m})-\nu_j(\boldsymbol{m})) + \mathscr{R}^{1,0}_{r_0,\infty}(|z_j|^2) + H) B_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle \bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} + \text{c.c.}\big) + (*).
$$

Then we reduce to the following system:

$$
\begin{aligned}
b_{\boldsymbol{m}}(\varrho(z)) &= \frac{\mathrm{i}}{\tilde{\boldsymbol{e}}(z)\cdot(\mu(\boldsymbol{m})-\nu(\boldsymbol{m}))}\big[a^{(\iota)}_{\boldsymbol{m}}(\varrho(z)) + \alpha_{\boldsymbol{m}}\big((b_{\boldsymbol{n}}(\varrho(z)))_{\boldsymbol{n}}, (B_{j\boldsymbol{n}}(\varrho_j(z)))_{j\boldsymbol{n}}, \varrho(z)\big)\big],\\
B_{j\boldsymbol{m}}(|z_j|^2) &= \mathrm{i}R_H\big(\boldsymbol{e}\cdot(\mu_j(\boldsymbol{m})-\nu_j(\boldsymbol{m})) + \mathscr{R}^{1,0}_{r_0,\infty}(|z_j|^2)\big)\\
&\qquad \times \big[G^{(\iota)}_{j\boldsymbol{m}}(|z_j|^2) + \Gamma_{j\boldsymbol{m}}\big(b(0,\dots,|z_j|^2,0,\dots,0), B(0,\dots,|z_j|^2,0,\dots,0), |z_j|^2\big)\big].
\end{aligned}
\tag{5-50}
$$

The $b_{\boldsymbol{m}}(\varrho(z))$ and $B_{j\boldsymbol{m}}(|z_j|^2)$ can be found by the implicit function theorem for $|z| < \delta'_\iota$ for δ'_ι sufficiently small. This gives us the desired polynomial χ, yielding $H^{(\iota+1)}$. Formulas (5-31) for the flow ϕ^t of χ are obtained choosing $\delta_\iota > 0$ sufficiently small, by Lemma 5.4(1). For the composition $\mathscr{F}^{(2N+4)}$, we obtain (5-34) as a consequence of (5-31) and of (4-44). □

6. Dispersion

We apply Theorem 5.9, set $\mathscr{H} = H^{(2N+4)}$, so that

$$
\mathscr{H}(z, \eta) = H_2(z, \eta) + \sum_{j=1}^{n} \lambda_j(|z_j|^2) + Z^{(2N+4)}(z, \boldsymbol{Z}, \eta) + \mathscr{R},
\tag{6-1}
$$

where

$$
\begin{aligned}
\mathscr{R} &:= \mathscr{R}^{1,2}_{r,\infty}(z, \eta) + \mathscr{R}^{0,2N+5}_{r,\infty}(z, \boldsymbol{Z}, \eta) + \mathrm{Re}\langle \boldsymbol{S}^{0,2N+4}_{r,\infty}(z, \boldsymbol{Z}, \eta), \bar{\eta}\rangle\\
&\quad + \sum_{i+j=2}\sum_{|\boldsymbol{m}|\le 1} \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{2\boldsymbol{m}ij}(z, \eta), \eta^i \bar{\eta}^j\rangle + \sum_{d+c=3}\sum_{i+j=d} \langle G_{dij}(z, \eta), \eta^i \bar{\eta}^j\rangle \mathscr{R}^{0,c}_{r,\infty}(z, \eta) + E_P(\eta).
\end{aligned}
\tag{6-2}
$$

Using formula (5-33) for $\iota = 2N+4$, we have

$$\sum_{j=1}^{n} \lambda_j(|z_j|^2) + Z^{(2N+4)}(z, \boldsymbol{Z}, \eta) = Z_0(z) + \sum_{j=1}^{n}\Bigg(\sum_{\boldsymbol{m}\in\mathcal{M}_j(2N+3)} \bar{z}_j \boldsymbol{Z}^{\boldsymbol{m}} \langle G_{j\boldsymbol{m}}(|z_j|^2), \eta\rangle + \text{c.c.}\Bigg) \tag{6-3}$$

$$\begin{aligned}\text{with}\quad Z_0(z) &:= \sum_{j=1}^{n} \lambda_j(|z_j|^2) + \sum_{\boldsymbol{m}\in\mathcal{M}_0(2N+4)} \boldsymbol{Z}^{\boldsymbol{m}} a_{\boldsymbol{m}}(|z_1|^2, \dots, |z_n|^2) \\ &= \mathscr{Z}_0(|z_1|^2, \dots, |z_n|^2),\end{aligned}$$

where the last equality holds for some $\mathscr{Z}_0(|z_1|^2, \dots, |z_n|^2)$ by Lemma 2.6.

Theorem 6.1 (main estimates). *There exist $\epsilon_0 > 0$ and $C_0 > 0$ such that, if the constant $0 < \epsilon$ of Theorem 1.3 satisfies $\epsilon < \epsilon_0$, then for $I = [0, \infty)$ and $C = C_0$ we have*

$$\|\eta\|_{L_t^p(I, W_x^{1,q})} \le C\epsilon \quad \textit{for all admissible pairs } (p, q), \tag{6-4}$$

$$\|z_j \boldsymbol{Z}^{\boldsymbol{m}}\|_{L_t^2(I)} \le C\epsilon \quad \textit{for all } (j, \boldsymbol{m}) \textit{ with } \boldsymbol{m} \in \mathcal{M}_j(2N+4), \tag{6-5}$$

$$\|z_j\|_{W_t^{1,\infty}(I)} \le C\epsilon \quad \textit{for all } j \in \{1, \dots, n\}. \tag{6-6}$$

Furthermore, there exists $\rho_+ \in [0, \infty)^n$ and a j_0 with $\rho_{+j} = 0$ for $j \neq j_0$, and an $\eta_+ \in H^1$ such that $\big|\rho_+ - |z(0)|\big| \le C\epsilon$ and $\|\eta_+\|_{H^1} \le C\epsilon$, such that

$$\lim_{t\to+\infty} \|\eta(t, x) - e^{\mathrm{i}t\Delta}\eta_+(x)\|_{H_x^1} = 0, \qquad \lim_{t\to+\infty} |z_j(t)| = \rho_{+j}. \tag{6-7}$$

Proof that Theorem 6.1 implies Theorem 1.3. Denote by (z', η') the initial coordinate system. By (5-35),

$$z' = z + \mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) \quad \text{and} \quad \eta' = \eta + \boldsymbol{S}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta).$$

Notice that (6-7) implies $\lim_{t\to+\infty} \boldsymbol{Z}(t) = 0$, and by standard arguments for $s > \frac{3}{2}$ we have

$$\lim_{t\to+\infty} \|e^{t\Delta}\eta_+\|_{L^{2,-s}(\mathbb{R}^3)} = 0 \quad \text{for any } \eta_+ \in L^2. \tag{6-8}$$

These two limits, Definitions 2.8–2.9 and (6-7) imply

$$\lim_{t\to+\infty} \mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = 0 \quad \text{in } \mathbb{C}^n \qquad \text{and} \qquad \lim_{t\to+\infty} \boldsymbol{S}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = 0 \quad \text{in } \Sigma_r.$$

This means that

$$\lim_{t\to+\infty} \|\eta'(t, x) - e^{\mathrm{i}t\Delta}\eta_+(x)\|_{H_x^1} = 0 \quad \text{and} \quad \lim_{t\to+\infty} |z'_j(t)| = \rho_{+j}, \tag{6-9}$$

so that (1-8) is true. Notice also that if we set $\tilde{\eta} = \eta$ and $A(t, x) = \boldsymbol{S}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta)$, we obtain the desired decomposition of η' satisfying (1-9) and (1-10). Finally, we have

$$\dot{z}'_j + \mathrm{i}e_j z'_j = \dot{z}_j + \mathrm{i}e_j z_j + \frac{d}{dt}\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) + \mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = O(\epsilon^2),$$

where $\dot z_j + \mathrm{i}e_j z_j = O(\epsilon^2)$ by (6-27) below, $\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = O(\epsilon^2)$ by (2-23) and $d\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta)/dt = O(\epsilon^2)$. To check the last of these, we write (it is easy that $d_w\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = \mathcal{R}^{1,0}_{r,\infty}(z, \boldsymbol{Z}, \eta)$ for $w = z, \boldsymbol{Z}$)

$$\frac{d}{dt}\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta) = \mathcal{R}^{1,0}_{r,\infty}(z, \boldsymbol{Z}, \eta)\dot z + \mathcal{R}^{1,0}_{r,\infty}(z, \boldsymbol{Z}, \eta)\dot{\boldsymbol{Z}} + d_\eta\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta)\cdot\dot\eta,$$

with $d_\eta\mathcal{R}^{1,1}_{r,\infty}$ the partial derivative in η. By a simple use of Taylor expansions and Definition 2.8,

$$\|d_\eta\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta)\|_{\Sigma^c_{-r}\to\Sigma^c_r} \le C(|z| + \|\eta\|_{\Sigma_{-r}}).$$

Then, by equations (6-12) and (6-27) below, we have $d\mathcal{R}^{1,1}_{r,\infty}(z, \boldsymbol{Z}, \eta)/dt = O(\epsilon^2)$. This yields the second inequality claimed in (1-9). $\square$

By a standard argument, (6-4)–(6-6) for $I = [0, \infty)$ are a consequence of the following proposition:

Proposition 6.2. *There exists a constant $c_0 > 0$ such that, for any $C_0 > c_0$, there is a value $\epsilon_0 = \epsilon_0(C_0)$ such that, if the inequalities (6-4)–(6-6) hold for $I = [0, T]$ for some $T > 0$, for $C = C_0$ and for $0 < \epsilon < \epsilon_0$, then, in fact, for $I = [0, T]$ the inequalities (6-4)–(6-6) hold for $C = C_0/2$.*

Proof. We will proceed via a series of lemmas.

Lemma 6.3. *Assume the hypotheses of Proposition 6.2 and take the M of Definition 2.5. Then there is a fixed c such that*

$$\|\eta\|_{L^p_t([0,T],W^{1,q})} \le c\epsilon + c\sum_{(\mu,\nu)\in M}|z^\mu\bar z^\nu|_{L^2_t(0,T)} \quad \text{for all admissible pairs } (p, q). \tag{6-10}$$

Proof. First of all, for $|z| < \delta_f$ and $\|\eta\|_{H^1\cap\mathcal{H}_c[0]} < \delta_f$ defining the domain of the Hamiltonian $\mathcal{H}(z, \eta)$ in (6-1), we will pick $\epsilon_0 \in (0, \delta_f)$ sufficiently small. Let $\epsilon \in (0, \epsilon_0)$, where $\epsilon = \|u(0)\|_{H^1}$. By (2-11), we have $|z'(0)| + \|\eta'(0)\|_X \le c_1\epsilon$, where $(z'(0), \eta'(0))$ are the coordinates in the initial system of coordinates introduced in Lemma 2.4. Let $(z(0), \eta(0))$ be the corresponding coordinates in the final system of coordinates. Then, by the relation (5-35), if ϵ_0 is sufficiently small we conclude that

$$|z(0)| + \|\eta(0)\|_{H^1} \le c_1'\epsilon \tag{6-11}$$

for some other fixed constant c_1'. We now turn to the equation of η. We have, for $\bar G_{j\boldsymbol{m}} = \bar G_{j\boldsymbol{m}}(0)$,

$$\mathrm{i}\dot\eta = \mathrm{i}\{\eta, \mathcal{H}\} = H\eta + \sum_{j=1}^{n}\sum_{l=1}^{2N+3}\sum_{|\boldsymbol{m}|=l} z_j\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\bar G_{j\boldsymbol{m}} + \mathbb{A}, \tag{6-12}$$

where

$$\mathbb{A} := \sum_{j=1}^{n}\sum_{l=1}^{2N+3}\sum_{|\boldsymbol{m}|=l} z_j\bar{\boldsymbol{Z}}^{\boldsymbol{m}}[\bar G_{j\boldsymbol{m}}(|z_j|^2) - \bar G_{j\boldsymbol{m}}] + \nabla_{\bar\eta}\mathcal{R}.$$

We rewrite

$$\sum_{j=1}^{n}\sum_{l=1}^{2N+3}\sum_{|\boldsymbol{m}|=l} z_j\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\bar G_{j\boldsymbol{m}} = \sum_{(\mu,\nu)\in M}\bar z^\mu z^\nu\bar G_{\mu\nu}. \tag{6-13}$$

Notice that (6-5) is the same as

$$\|z^{\mu}\bar{z}^{\nu}\|_{L^2_t(I)} \le C\epsilon \quad \text{for all } (\mu,\nu)\in M. \tag{6-14}$$

Suppose we can show that, for $I_T := [0,T]$,

$$\|\mathbb{A}\|_{L^2(I_T,H^{1,S})+L^1(I_T,H^1)} \le C(S,C_0)\epsilon^2. \tag{6-15}$$

Then, if ϵ_0 is small enough and $\epsilon\in(0,\epsilon_0)$, we obtain (6-10) by $H^{1,S}(\mathbb{R}^3)\hookrightarrow W^{1,6/5}(\mathbb{R}^3)$, by (6-11), (6-14) and (6-15) and by the Strichartz estimates, which, for P_c the orthogonal projection of L^2 onto $\mathcal{H}[0]$, are valid for P_cH by [Yajima 1995] (here notice that all the terms in (6-12) belong to $\mathcal{H}[0]$).

So now we prove (6-15). We have, for $r-1\ge S>\frac{9}{2}$,

$$\begin{aligned}\|z_j\bar{\boldsymbol{Z}}^{\boldsymbol{m}}[\bar{G}_{j\boldsymbol{m}}(|z_j|^2)-\bar{G}_{j\boldsymbol{m}}]\|_{L^2(I_T,H^{1,S})} &\le \|z_j\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\|_{L^2(I_T,\mathbb{C})}\|\bar{G}_{j\boldsymbol{m}}(|z_j|^2)-\bar{G}_{j\boldsymbol{m}}\|_{L^\infty(I_T,H^{1,S})}\\ &\le C_0\epsilon\sup\{\|G'_{j\boldsymbol{m}}(|z_j|^2)\|_{\Sigma_r} : |z_j|\le\delta_0\}\|z_j^2\|_{L^\infty(I_T,\mathbb{C})}\\ &\le CC_0^3\epsilon^3<c\epsilon.\end{aligned} \tag{6-16}$$

We have, for a fixed $c_1>0$,

$$\|\nabla_\eta E_P(\eta)\|_{L^1(I_T,H^1)} = 2\||\eta|^2\eta\|_{L^1(I_T,H^1)} \le c_1\|\eta\|_{L^\infty(I_T,H^1)}\|\eta\|^2_{L^2(I_T,L^6)} \le c_1C_0^3\epsilon^3. \tag{6-17}$$

We finally show that, for an arbitrarily preassigned $S>2$,

$$\|R_1\|_{L^2(I_T,H^{1,S})} \le C(S,C_0)\epsilon^2 \quad \text{for } R_1=\nabla_\eta(\mathcal{R}-E_P(\eta)). \tag{6-18}$$

R_1 is a sum of various terms obtained from the expansion (6-2). Let us start by showing

$$\|\nabla_{\bar\eta}\mathcal{R}^{1,2}_{r,\infty}(z,\eta)\|_{L^2(I_T,H^{1,S})} \le C(S,C_0)\epsilon^2. \tag{6-19}$$

Recalling (2-25), it is elementary to show that $\nabla_{\bar\eta}\mathcal{R}^{1,2}_{r,\infty}(z,\eta)=\boldsymbol{S}^{1,1}_{r,\infty}(z,\eta)$ and

$$\begin{aligned}\|\boldsymbol{S}^{1,1}_{r,\infty}(z,\eta)\|_{L^2(I_T,H^{1,S})} &\le C_1\|(\|\eta\|_{\Sigma_{-r}}+|z|)\|_{L^\infty(I_T)}\|\eta\|_{L^2(I_T,\Sigma_{-r})}\\ &\le C_2\|(\|\eta\|_{H^1}+|z|)\|_{L^\infty(I_T)}\|\eta\|_{L^2(I_T,L^6)} \le C(S,C_0)\epsilon^2.\end{aligned}$$

We next show

$$\|\nabla_{\bar\eta}\mathcal{R}^{0,2N+5}_{r,\infty}(z,\boldsymbol{Z},\eta)\|_{L^2(I_T,H^{1,S})} \le C(S,C_0)\epsilon^2. \tag{6-20}$$

We have, for a remainder $\|O(\|\eta\|^2_{\Sigma_{-r}})\|_{\Sigma_r}\le C\|\eta\|^2_{\Sigma_{-r}}$,

$$\nabla_{\bar\eta}\mathcal{R}^{0,2N+5}_{r,\infty}(z,\boldsymbol{Z},\eta)=\boldsymbol{S}^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z},\eta)=\boldsymbol{S}^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z})+d_\eta\boldsymbol{S}^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z},0)\cdot\eta+O(\|\eta\|^2_{\Sigma_{-r}}).$$

We have, by Lemma 2.7,

$$\begin{aligned}\|\boldsymbol{S}^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z})\|_{L^2(I_T,H^{1,S})} &\le C_1\sup_{|z|\le C_0\epsilon}\|\boldsymbol{S}^{0,0}_{r,\infty}(z,\boldsymbol{Z})\|_{\Sigma_{M'}}\||\boldsymbol{Z}|^{2N+4}\|_{L^2(I_T)}\\ &\le C_2\|z\|_{L^\infty(I)}\sum_j\sum_{(\mu,\nu)\in M_j(N+1)}\|z^\mu\bar{z}^\nu\|_{L^\infty(I_T)}\|z^\mu\bar{z}^\nu\|_{L^2(I_T)}\\ &\le C(S,C_0)\epsilon^3.\end{aligned}$$

We have

$$\begin{aligned}\|d_\eta S^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z},0)\cdot\eta\|_{L^2(I_T,H^{1,S})} &\le C_1(S)\|\eta\|_{L^2(I_T,\Sigma_{-r})}\sup_{|z|\le C_0\epsilon}\|d_\eta S^{0,2N+4}_{r,\infty}(z,\boldsymbol{Z},0)\|_{\Sigma_{-r}\to\Sigma_r}\\ &\le C_2(S)\|\eta\|_{L^2(I_T,L^6)}\sup_{|z|\le C_0\epsilon}|\boldsymbol{Z}|^{2N+3}\\ &\le C(S,C_0)\epsilon^2.\end{aligned}$$

Hence (6-20) is proved. Other terms in R_1 can be bounded with similarly elementary arguments, yielding (6-18). Then (6-16), (6-17) and (6-18) imply (6-15). □

Setting $M=M(2N+4)$ — see Definition 2.5 — we now introduce a new variable g, setting

$$g=\eta+Y \quad \text{with} \quad Y:=\sum_{(\alpha,\beta)\in M}\bar{z}^\alpha z^\beta R^+_H(\boldsymbol{e}\cdot(\beta-\alpha))\bar{G}_{\alpha\beta}. \tag{6-21}$$

Lemma 6.4. *Assume the hypotheses of Proposition 6.2 and fix $S>\frac{9}{2}$. Then there is a $c_1(S)>0$ such that, for any C_0, there is an $\epsilon_0=\epsilon_0(C_0,S)>0$ such that, for $\epsilon\in(0,\epsilon_0)$ in Theorem 1.3, we have*

$$\|g\|_{L^2([0,T],L^{2,-S})}\le c_1(S)\epsilon. \tag{6-22}$$

Proof. We have

$$\mathrm{i}\dot{g}=Hg+\mathbb{A}+\boldsymbol{T}, \quad \text{where} \quad \boldsymbol{T}:=\sum_j[\partial_{z_j}Y(\mathrm{i}\dot{z}_j-e_jz_j)+\partial_{\bar{z}_j}Y(\mathrm{i}\dot{\bar{z}}_j+e_j\bar{z}_j)]. \tag{6-23}$$

We then have

$$g(t)=e^{-\mathrm{i}Ht}\eta(0)+e^{-\mathrm{i}Ht}Y(0)-\mathrm{i}\int_0^t e^{-\mathrm{i}H(t-s)}(\mathbb{A}(s)+\boldsymbol{T}(s))\,ds. \tag{6-24}$$

We have, for fixed constants, by (6-11) and (6-15), the inequalities

$$\|e^{-\mathrm{i}Ht}\eta(0)\|_{L^2([0,T],L^{2,-S})}a\le c_2\|e^{-\mathrm{i}Ht}\eta(0)\|_{L^2([0,T],L^6)}\le c_2'\|\eta(0)\|_{L^2}\le c_3\epsilon,$$

$$\left\|\int_0^t e^{-\mathrm{i}H(t-s)}\mathbb{A}(s)\,ds\right\|_{L^2([0,T],L^{2,-S})}\le c_2\|\mathbb{A}\|_{L^2([0,T],H^{1,S})+L^1([0,T],H^1)}\le C(C_0,S)\epsilon^2.$$

For a proof of the following standard lemma see, for instance, the proof of [Cuccagna 2003, Lemma 5.4].

Lemma 6.5. *Let Λ be a compact subset of $(0,\infty)$ and let $S>\frac{9}{2}$. Then there exists a fixed $c(S,\Lambda)$ such that, for every $t\ge 0$ and $\lambda\in\Lambda$,*

$$\|e^{-\mathrm{i}Ht}R^+_H(\lambda)P_cv_0\|_{L^{2,-S}(\mathbb{R}^3)}\le c(S,\Lambda)\langle t\rangle^{-3/2}\|P_cv_0\|_{L^{2,S}(\mathbb{R}^3)} \quad \textit{for all } v_0\in L^{2,S}(\mathbb{R}^3).$$

By Lemma 6.5, (6-11) and $G_{\alpha\beta}=P_cG_{\alpha\beta}$, we have

$$\begin{aligned}\|e^{-\mathrm{i}Ht}Y(0)\|_{L^2([0,T],L^{2,-S})}&\le\sum_{(\alpha,\beta)\in M}|z^\alpha(0)z^\beta(0)|\,\|e^{-\mathrm{i}Ht}R^+_H(\boldsymbol{e}\cdot(\beta-\alpha))\bar{G}_{\alpha\beta}\|_{L^2([0,T],L^{2,-S})}\\ &\le(\sharp M)c_2\epsilon^2\|\langle t\rangle^{-3/2}\|_{L^2(0,T)}c(S,\Lambda)\|\bar{G}_{\alpha\beta}\|_{L^{2,S}}\le C(N,C_0,S)\epsilon^2\end{aligned}$$

with $\sharp M$ the cardinality of M and a fixed c_2, and where the set Λ is as in Lemma 6.5,

$$\Lambda := \{(\nu-\mu)\cdot \boldsymbol{e} : (\mu,\nu)\in M\}. \tag{6-25}$$

We finally consider, for definiteness (the term $\partial_{\bar z_j} Y(\mathrm{i}\dot{\bar z}_j + e_j \bar z_j)$ can be treated similarly),

$$\begin{aligned}
&\left\| \int_0^t e^{-\mathrm{i}H(t-s)} R_H^+(\boldsymbol{e}\cdot(\beta-\alpha))\overline{G}_{\alpha\beta}\partial_{z_j} Y(s)(\mathrm{i}\dot z_j - e_j z_j)(s)\,ds \right\|_{L^2([0,T],L^{2,-S})}\\
&\qquad \le c(S,\Lambda)\sum_{(\alpha,\beta)\in M}\|G_{\alpha\beta}\|_{L^{2,S}}\beta_j \left\| \int_0^t \langle t-s\rangle^{-3/2}\left| \frac{\bar z^\alpha(s) z^\beta(s)}{z_j(s)}(\mathrm{i}\dot z_j - e_j z_j)(s)\right| ds\right\|_{L^2(0,T)}\\
&\qquad \le c(S,\Lambda)c_2 \sum_{(\alpha,\beta)\in M}\beta_j \left\| \frac{\bar z^\alpha(s) z^\beta}{z_j}(\mathrm{i}\dot z_j - e_j z_j)\right\|_{L^2(0,T)}
\end{aligned} \tag{6-26}$$

for fixed c_2. We have

$$\begin{aligned}
\mathrm{i}\dot z_j &= (1+\varpi_j(|z_j|^2))(e_j z_j + \partial_{\bar z_j}\mathscr{L}_0(|z_1|^2,\dots,|z_n|^2) + \partial_{\bar z_j}\mathscr{R})\\
&\quad + (1+\varpi_j(|z_j|^2))\Bigg[\sum_{(\mu,\nu)\in M} \nu_j \frac{z^\mu \bar z^\nu}{\bar z_j}\langle \eta, G_{\mu\nu}\rangle + \sum_{(\mu',\nu')\in M}\mu'_j \frac{z^{\nu'}\bar z^{\mu'}}{\bar z_j}\langle \bar\eta, \overline{G}_{\mu'\nu'}\rangle \Bigg]\\
&\qquad + (1+\varpi_j(|z_j|^2))\Bigg[\sum_{\boldsymbol{m}\in\mathcal{M}_j(2N+3)} |z_j|^2 \boldsymbol{Z}^{\boldsymbol{m}}\langle G'_{j\boldsymbol{m}},\eta\rangle + z_j^2 \overline{\boldsymbol{Z}}^{\boldsymbol{m}}\langle \overline{G}'_{j\boldsymbol{m}},\bar\eta\rangle\Bigg].
\end{aligned} \tag{6-27}$$

To bound (6-26), we substitute $(\mathrm{i}\dot z_j - e_j z_j)$ by the other terms in (6-27) in the last line of (6-26) . So, for example, we have $\partial_{\bar z_j}\mathscr{L}_0(|z_1|^2,\dots,|z_n|^2) \sim z_j O(\epsilon)$, which by (6-14) yields

$$\beta_j \left\| \frac{\bar z^\alpha z^\beta}{z_j}\partial_{\bar z_j}\mathscr{L}_0(|z_1|^2,\dots,|z_n|^2)\right\|_{L^2(0,T)} \le C(C_0)\epsilon \|\bar z^\alpha z^\beta\|_{L^2(0,T)} \le C(C_0)C_0\epsilon^2.$$

For $(\mu,\nu)\in M$, we have, in $(0,T)$,

$$\beta_j\nu_j \left\| \frac{\bar z^\alpha z^\beta}{z_j}\frac{z^\mu \bar z^\nu}{\bar z_j}\langle \eta, G_{\mu\nu}\rangle\right\|_{L^2_t} \le \beta_j\nu_j \left\| \frac{\bar z^\alpha z^\beta}{z_j}\frac{z^\mu \bar z^\nu}{\bar z_j}\right\|_{L^\infty_t} \|G_{\mu\nu}\|_{L^{\frac{6}{5}}}\|\eta\|_{L^\infty_t L^6} \le C(C_0)\epsilon^2.$$

A similar argument works for the terms in the second summation in the second line of (6-27). Finally,

$$\beta_j \left\| \frac{\bar z^\alpha z^\beta}{z_j}\partial_{\bar z_j}\mathscr{R}\right\|_{L^2(0,T)} \le \beta_j \left\| \frac{\bar z^\alpha z^\beta}{z_j}\right\|_{L^\infty(0,T)} \|\partial_{\bar z_j}\mathscr{R}\|_{L^2(0,T)} \le C(C_0)\epsilon^3$$

is a consequence of the bound

$$\|\partial_{\bar z_j}\mathscr{R}\|_{L^p(0,T)} \le C(C_0)\epsilon^2 \quad \text{for any } p\in[1,\infty]. \tag{6-28}$$

Here we need to check (6-28) term by term for the sum in the right-hand side of (6-2). This is straightforward using (2-23), (2-25) and (2-26) and the fact, stated in Lemma 5.8, that $G_{2\boldsymbol{m}ij}$ and G_{dij} are $\boldsymbol{S}^{0,0}_{r,\infty}$. $\square$

We turn now to the Fermi golden rule (FGR). We substitute (6-21) into (6-27), getting

$$\mathrm{i}\dot z_j=(1+\varpi_j(|z_j|^2))(e_jz_j+\partial_{\bar z_j}\mathscr{Z}_0(|z_1|^2,\dots,|z_n|^2))-\sum_{\substack{(\mu,\nu)\in M\\(\alpha,\beta)\in M}}\nu_j\frac{z^{\mu+\beta}\bar z^{\nu+\alpha}}{\bar z_j}\langle R_H^+(\boldsymbol{e}\cdot(\beta-\alpha))\bar G_{\alpha\beta},G_{\mu\nu}\rangle$$
$$-\sum_{\substack{(\mu',\nu')\in M\\(\alpha',\beta')\in M}}\mu'_j\frac{z^{\nu'+\alpha'}\bar z^{\mu'+\beta'}}{\bar z_j}\langle R_H^-(\boldsymbol{e}\cdot(\beta'-\alpha'))G_{\alpha'\beta'},\bar G_{\mu'\nu'}\rangle+\mathscr{F}_j,\tag{6-29}$$

where

$$\mathscr{F}_j:=(1+\varpi_j(|z_j|^2))\partial_{\bar z_j}\mathscr{R}+\varpi_j(|z_j|^2)\Bigg[\sum_{(\mu,\nu)\in M}\nu_j\frac{z^\mu\bar z^\nu}{\bar z_j}\langle\eta,G_{\mu\nu}\rangle+\sum_{(\mu',\nu')\in M}\mu'_j\frac{z^{\nu'}\bar z^{\mu'}}{\bar z_j}\langle\bar\eta,\bar G_{\mu'\nu'}\rangle\Bigg]$$
$$+\sum_{(\mu,\nu)\in M}\nu_j\frac{z^\mu\bar z^\nu}{\bar z_j}\langle g,G_{\mu\nu}\rangle+\sum_{(\mu',\nu')\in M}\mu'_j\frac{z^{\nu'}\bar z^{\mu'}}{\bar z_j}\langle\bar g,\bar G_{\mu'\nu'}\rangle$$
$$+(1+\varpi_j(|z_j|^2))\Bigg[\sum_{\boldsymbol{m}\in\mathscr{M}_j(2N+3)}|z_j|^2\boldsymbol{Z}^{\boldsymbol{m}}\langle G'_{j\boldsymbol{m}},\eta\rangle+z_j^2\bar{\boldsymbol{Z}}^{\boldsymbol{m}}\langle\bar G'_{j\boldsymbol{m}},\bar\eta\rangle\Bigg].\tag{6-30}$$

We introduce the new variable ζ, defined by

$$z_j-\zeta_j=-\sum_{\substack{(\mu,\nu)\in M\\(\alpha,\beta)\in M}}\frac{\nu_jz^{\mu+\beta}\bar z^{\nu+\alpha}}{((\mu-\nu)\cdot\boldsymbol{e}-(\alpha-\beta)\cdot\boldsymbol{e})\bar z_j}\langle R_H^+(\boldsymbol{e}\cdot(\beta-\alpha))\bar G_{\alpha\beta},G_{\mu\nu}\rangle$$
$$-\sum_{\substack{(\mu',\nu')\in M\\(\alpha',\beta')\in M}}\frac{\mu'_jz^{\nu'+\alpha'}\bar z^{\mu'+\beta'}}{((\alpha'-\beta')\cdot\boldsymbol{e}-(\mu'-\nu')\cdot\boldsymbol{e})\bar z_j}\langle R_H^-(\boldsymbol{e}\cdot(\beta'-\alpha'))G_{\alpha'\beta'},\bar G_{\mu'\nu'}\rangle,\tag{6-31}$$

where we are summing only on pairs where the formula makes sense (i.e., only on pairs not in the same set M_L for an $L\in\Lambda$; see (6-33) below). It is easy to see that

$$\|\zeta-z\|_{L^2(0,T)}\le c(N,C_0)\epsilon^2\quad\text{and}\quad\|\zeta-z\|_{L^\infty(0,T)}\le c(N,C_0)\epsilon^2.\tag{6-32}$$

Recall now the set $\Lambda=\{(\nu-\mu)\cdot\boldsymbol{e}:(\mu,\nu)\in M\}$ defined in (6-25). For any $L\in\Lambda$, set

$$M_L:=\{(\mu,\nu)\in M:(\nu-\mu)\cdot\boldsymbol{e}=L\}.\tag{6-33}$$

We then get

$$\mathrm{i}\dot\zeta_j=(1+\varpi(|z_j|^2))(e_j\zeta_j+\partial_{\bar j}\mathscr{Z}_0(|\zeta_1|^2,\dots,|\zeta_n|^2))-\sum_{L\in\Lambda}\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\nu_j\frac{\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}}{\bar\zeta_j}\langle R_H^+(\boldsymbol{e}\cdot(\beta-\alpha))\bar G_{\alpha\beta},G_{\mu\nu}\rangle$$
$$-\sum_{L\in\Lambda}\sum_{\substack{(\mu',\nu')\in M_L\\(\alpha',\beta')\in M_L}}\mu'_j\frac{\zeta^{\nu'+\alpha'}\bar\zeta^{\mu'+\beta'}}{\bar\zeta_j}\langle R_H^-(\boldsymbol{e}\cdot(\beta'-\alpha'))G_{\alpha'\beta'},\bar G_{\mu'\nu'}\rangle+\mathscr{G}_j,\tag{6-34}$$

where, for some $A_{k\alpha\beta\mu\nu}$, $B_{k\alpha\beta\mu\nu}$, we have

$$\begin{aligned}\mathcal{G}_j &= \mathcal{F}_j + (1+\varpi(|z_j|^2))[\partial_{\bar{j}}\mathcal{Z}_0(|z_1|^2,\dots,|z_n|^2) - \partial_{\bar{j}}\mathcal{Z}_0(|\zeta_1|^2,\dots,|\zeta_n|^2)] \\ &\quad - e_j\varpi(|z_j|^2)\Bigg[\sum_{\substack{(\mu,\nu)\in M\\(\alpha,\beta)\in M}} \frac{\nu_j z^{\mu+\beta}\bar{z}^{\nu+\alpha}}{((\mu-\nu)\cdot\boldsymbol{e}-(\alpha-\beta)\cdot\boldsymbol{e})\bar{z}_j}\langle R_H^+(\boldsymbol{e}\cdot(\beta-\alpha))\bar{G}_{\alpha\beta}, G_{\mu\nu}\rangle \\ &\qquad + \sum_{\substack{(\mu',\nu')\in M\\(\alpha',\beta')\in M}} \frac{\mu'_j z^{\nu'+\alpha'}\bar{z}^{\mu'+\beta'}}{((\alpha'-\beta')\cdot\boldsymbol{e}-(\mu'-\nu')\cdot\boldsymbol{e})\bar{z}_j}\langle R_H^-(\boldsymbol{e}\cdot(\beta'-\alpha'))G_{\alpha'\beta'}, \bar{G}_{\mu'\nu'}\rangle\Bigg] \\ &\quad + \sum_k \sum_{\substack{(\mu,\nu)\in M\\(\alpha,\beta)\in M}} (\mathrm{i}\dot{z}_k - e_k z_k)\frac{z^{\mu+\beta}\bar{z}^{\nu+\alpha}}{\bar{z}_j}A_{k\alpha\beta\mu\nu} + \overline{\mathrm{i}\dot{z}_k - e_k z_k}\,\frac{z^{\mu+\beta}\bar{z}^{\nu+\alpha}}{\bar{z}_j}B_{k\alpha\beta\mu\nu}. \end{aligned} \tag{6-35}$$

Lemma 6.6. *There are fixed c_4 and $\epsilon_0 > 0$ such that, for $\epsilon \in (0,\epsilon_0)$, we have*

$$\|\mathcal{G}_j\bar{\zeta}_j\|_{L^1[0,T]} \le (1+C_0)c_4\epsilon^2. \tag{6-36}$$

Proof. We consider separately the terms in the right-hand side of (6-35) and (6-30). By (6-6), (6-28) and (6-32),

$$\|\partial_{\bar{z}_j}\mathcal{R}\bar{\zeta}_j\|_{L^1_t[0,T]} \le C(C_0)\epsilon^3.$$

For fixed constants c_2 and c_3, by (6-4) and (6-22) we have

$$\left\|\frac{z^\mu\bar{z}^\nu\bar{\zeta}_j}{\bar{z}_j}\langle g, G_{\mu\nu}\rangle\right\|_{L^1[0,T]} \le c_2\left\|\frac{z^\mu\bar{z}^\nu\bar{\zeta}_j}{\bar{z}_j}\right\|_{L^2[0,T]}\|g\|_{L^2([0,T],L^{2,-S})} \le c_3C_0\epsilon^2. \tag{6-37}$$

To get (6-37) we exploit Lemma 6.4 and the following bound:

$$\nu_j\left\|\frac{z^\mu\bar{z}^\nu\bar{\zeta}_j}{\bar{z}_j}\right\|_{L^2[0,T]} \le \nu_j\|z^\mu\bar{z}^\nu\|_{L^2[0,T]} + \nu_j\left\|\frac{z^\mu\bar{z}^\nu}{\bar{z}_j}\right\|_{L^\infty[0,T]}\|\zeta_j-\bar{z}_j\|_{L^2[0,T]} \le c_2C_0\epsilon + C(C_0)\epsilon^3 \tag{6-38}$$

for fixed c_2, where we used (6-14) and (6-32). Terms such as (6-37), that is, the terms from the second term in the right-hand side of (6-30), are the ones responsible for the $C_0c_4\epsilon^2$ in (6-36), where C_0 could be large. The other terms are $O(\epsilon^2)$ with fixed constants if ϵ_0 is small enough.

By (6-4) and (6-5), for $\boldsymbol{m} \in \mathcal{M}_j(2N+4)$ we have

$$\|\,|z_j|^2\boldsymbol{Z}^{\boldsymbol{m}}\langle G'_{j\boldsymbol{m}}, \eta\rangle\bar{\zeta}_j\|_{L^1[0,T]} \le c_4\|z_j\zeta_j\|_{L^\infty}\|z_j\boldsymbol{Z}^{\boldsymbol{m}}\|_{L^2[0,T]}\|\eta\|_{L^2([0,T],L^{2,-S})} \le C(C_0)\epsilon^4. \tag{6-39}$$

Let $\mathfrak{S}$ be the sum of the second to fourth lines in (6-35). It is easy to see by (6-32) that

$$\|\bar{\zeta}_j(\mathfrak{S})\|_{L^2[0,T]} \le C(C_0)\epsilon^3; \tag{6-40}$$

see [Cuccagna 2011b, Lemma 4.11]. Furthermore,

$$\|[\partial_{\bar{j}}\mathcal{Z}_0(|z_1|^2,\dots,|z_n|^2) - \partial_{\bar{j}}\mathcal{Z}_0(|\zeta_1|^2,\dots,|\zeta_n|^2)]\bar{\zeta}_j\|_{L^2[0,T]} \le C(C_0)\epsilon^3; \tag{6-41}$$

see [Cuccagna 2011b, Lemma 4.10]. Finally we have, for $(\mu,\nu)\in M$,

$$\begin{aligned}\Big\|\varpi_j(|z_j|^2)\nu_j\frac{z^\mu\bar z^\nu}{\bar z_j}\langle\eta,G_{\mu\nu}\rangle\zeta_j\Big\|_{L^1_t}&\\ \le \|\varpi_j(|z_j|^2)\nu_j z^\mu\bar z^\nu\langle\eta,G_{\mu\nu}\rangle\|_{L^1_t}&+\Big\|\varpi_j(|z_j|^2)\nu_j\frac{z^\mu\bar z^\nu}{\bar z_j}\langle\eta,G_{\mu\nu}\rangle(\zeta_j-z_j)\Big\|_{L^1_t}\\ \le C(C_0)\epsilon^3&\end{aligned}$$

by $\varpi_j(|z_j|^2)=O(|z_j|^2)$, (6-4)–(6-6) and (6-32). This completes the proof of Lemma 6.6. ☐

We now consider

$$\begin{aligned}2^{-1}\frac{d}{dt}\sum_j|e_j|\,|\zeta_j|^2=-\sum_j e_j\,&\overbrace{\mathrm{Im}[(1+\varpi(|z_j|^2))e_j|\zeta_j|^2+\partial_{\bar\zeta_j}\mathscr{Z}_0(|\zeta_1|^2,\dots,|\zeta_n|^2)\bar\zeta_j]}^{0}\\ &-\sum_j e_j\,\mathrm{Im}[\mathscr{G}_j\bar\zeta_j]+\sum_{L\in\Lambda}\mathrm{Im}\Big[\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\nu\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\langle R^+_H(L)\bar G_{\alpha\beta},G_{\mu\nu}\rangle\\ &+\sum_{\substack{(\mu',\nu')\in M_L\\(\alpha',\beta')\in M_L}}\mu'\cdot\boldsymbol{e}\zeta^{\nu'+\alpha'}\bar\zeta^{\mu'+\beta'}\langle R^-_H(L)G_{\alpha'\beta'},\bar G_{\mu'\nu'}\rangle\Big].\end{aligned}\tag{6-42}$$

We can now substitute $R^\pm_H(L)=\mathrm{P.V.}(1/(H-L))\pm\mathrm{i}\pi\delta(H-L)$.

Lemma 6.7. *The contributions to* (6-42) *from* P.V.$(1/(H-L))$ *cancel out*:

$$\begin{aligned}\mathrm{Im}\Big[\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\nu\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\Big\langle\mathrm{P.V.}\frac{1}{H-L}\bar G_{\alpha\beta},G_{\mu\nu}\Big\rangle&\\ +\sum_{\substack{(\mu',\nu')\in M_L\\(\alpha',\beta')\in M_L}}\mu'\cdot\boldsymbol{e}\zeta^{\nu'+\alpha'}\bar\zeta^{\mu'+\beta'}\Big\langle\mathrm{P.V.}\frac{1}{H-L}G_{\alpha'\beta'},\bar G_{\mu'\nu'}\Big\rangle\Big]&=0.\end{aligned}\tag{6-43}$$

Proof. We set $(\alpha',\beta')=(\mu,\nu)$ and $(\mu',\nu')=(\alpha,\beta)$ in the second sum of (6-43). With these choices,

$$\mu'\cdot\boldsymbol{e}\zeta^{\nu'+\alpha'}\bar\zeta^{\mu'+\beta'}\Big\langle\mathrm{P.V.}\frac{1}{H-L}G_{\alpha'\beta'},\bar G_{\mu'\nu'}\Big\rangle=\alpha\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\Big\langle\mathrm{P.V.}\frac{1}{H-L}\bar G_{\alpha\beta},G_{\mu\nu}\Big\rangle.$$

Then 2 times the left-hand side of (6-43) becomes

$$\begin{aligned}&2\,\mathrm{Im}\Big[\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}(\alpha+\nu)\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\Big\langle\mathrm{P.V.}\frac{1}{H-L}\bar G_{\alpha\beta},G_{\mu\nu}\Big\rangle\Big]\\ &\quad=\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\mathrm{Im}\Big[(\alpha+\nu)\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\Big\langle\mathrm{P.V.}\frac{1}{H-L}\bar G_{\alpha\beta},G_{\mu\nu}\Big\rangle\\ &\qquad\qquad+(\mu+\beta)\cdot\boldsymbol{e}\bar\zeta^{\mu+\beta}\zeta^{\nu+\alpha}\Big\langle\mathrm{P.V.}\frac{1}{H-L}\bar G_{\mu\nu},G_{\alpha\beta}\Big\rangle\Big]\end{aligned}$$

$$= \operatorname{Im}\Big[\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}(\alpha+\nu)\cdot\boldsymbol{e}\Big(\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\Big\langle \text{P.V.}\frac{1}{H-L}\bar G_{\alpha\beta},G_{\mu\nu}\Big\rangle+\text{c.c.}\Big)\Big]=0,$$

where we exploited the fact that, if (μ,ν) and (α,β) both belong to M_L, then $(\alpha+\nu)\cdot\boldsymbol{e}=(\mu+\beta)\cdot\boldsymbol{e}$. $\square$

Lemma 6.8. *Set, for any $L\in\Lambda$,*

$$G_L(\zeta):=\sqrt{\pi}\sum_{(\mu,\nu)\in M_L}\zeta^\mu\bar\zeta^\nu G_{\mu\nu}. \tag{6-44}$$

Then we have

$$\operatorname{Im}\Big[\mathrm{i}\pi\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\nu\cdot\boldsymbol{e}\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\langle\delta(H-L)\bar G_{\alpha\beta},G_{\mu\nu}\rangle+\mathrm{i}\pi\sum_{\substack{(\mu',\nu')\in M_L\\(\alpha',\beta')\in M_L}}\mu'\cdot\boldsymbol{e}\zeta^{\nu'+\alpha'}\bar\zeta^{\mu'+\beta'}\langle\delta(H-L)G_{\alpha'\beta'},\bar G_{\mu'\nu'}\rangle\Big]$$
$$=L\langle\delta(H-L)\bar G_L(\zeta),G_L(\zeta)\rangle\ge 0. \tag{6-45}$$

Proof. First of all, the last inequality is a consequence of the formula

$$\langle F,\delta(H-L)\bar G\rangle=\frac{1}{2\sqrt{L}}\int_{|\xi|=\sqrt{L}}\widehat F(\xi)\overline{\widehat G}(\xi)\,d\sigma(\xi)$$

with $\widehat F$ and $\widehat G$ the Fourier transforms of F and G associated to H; see [Taylor 1997, Chapter 9, Proposition 2.2].

To prove the equality in (6-45), set $(\alpha',\beta')=(\alpha,\beta)$ and $(\mu',\nu')=(\mu,\nu)$ in the second sum of (6-45). Then the left-hand side of (6-45) equals

$$\pi\operatorname{Re}\Big[\sum_{\substack{(\mu,\nu)\in M_L\\(\alpha,\beta)\in M_L}}\overbrace{(\nu-\mu)\cdot\boldsymbol{e}}^{L}\,\zeta^{\mu+\beta}\bar\zeta^{\nu+\alpha}\langle\delta(H-L)\bar G_{\alpha\beta},G_{\mu\nu}\rangle\Big]=L\langle\delta(H-L)\bar G_L(\zeta),G_L(\zeta)\rangle. \qquad\square$$

From (6-42) and Lemmas 6.7–6.8, we obtain

$$2\sum_{L\in\Lambda}L\langle\delta(H-L)\bar G_L(\zeta),G_L(\zeta)\rangle=\frac{d}{dt}\sum_j|e_j||\zeta_j|^2+2\sum_j e_j\operatorname{Im}[\mathcal{G}_j\bar\zeta_j]. \tag{6-46}$$

We are able to restate, precisely this time, hypothesis (H4):

(H4) For some fixed constants, we have

$$\sum_{L\in\Lambda}\langle\delta(H-L)\bar G_L(\zeta),G_L(\zeta)\rangle\sim\sum_{(\mu,\nu)\in M}|\zeta^{\mu+\nu}|^2\quad\text{for all }\zeta\in\mathbb{C}^n\text{ with }|\zeta|\le 1. \tag{6-47}$$

We now complete the proof of Proposition 6.2. We claim we have, for a fixed c,

$$\Big|\sum_j|e_j|\big(|\zeta_j(t)|^2-|\zeta_j(0)|^2\big)\Big|\le c\epsilon^2. \tag{6-48}$$

Indeed, first of all we have $|\zeta_j(0)| \le c'\epsilon$ by $\epsilon := \|u_0\|_{H^1}$. Observe that, for (z', η') the initial coordinates in Lemma 2.4, by Proposition 1.1 and Lemma 2.3 it is easy to see that we have

$$\epsilon^2 > \|u_0\|_{L^2}^2 = \|u(t)\|_{L^2}^2 = \left\| \left(\sum_{j=1}^n z_j'(t)\phi_j + \eta'(t) \right) + \left(\sum_{j=1}^n q_{j z_j'(t)} + (R[z'(t)] - 1)\eta'(t) \right) \right\|_{L^2}^2$$
$$= \sum_{j=1}^n |z_j'(t)|^2 + \|\eta'(t)\|_{L^2}^2 + O\big(|z'(t)|^6 + |z'(t)|^4 \|\eta'(t)\|_{L^2}^2\big).$$

This gives the following version of (2-11):

$$\sum_{j=1}^n |z_j'(t)|^2 + \|\eta'(t)\|_{L^2}^2 \le 2\epsilon^2. \tag{6-49}$$

This yields an analogous formula for the last system of coordinates, (z, η) in (5-35). Finally, this yields the following inequality for the variables ζ introduced in (6-31):

$$\sum_{j=1}^n |\zeta_j(t)|^2 \le 3\epsilon^2. \tag{6-50}$$

Hence the claim (6-48) is proved. By Lemma 6.6, the hypothesis (6-47), (6-32) and (6-48), for $\epsilon \in (0, \epsilon_0)$ with ϵ_0 small enough we obtain, for a fixed c,

$$\sum_{(\mu,\nu)\in M} \|z^{\mu+\nu}\|_{L^2(0,t)}^2 \le c\epsilon^2 + cC_0\epsilon^2. \tag{6-51}$$

Now, (6-51) tells us that $\|z^{\mu+\nu}\|_{L^2(0,t)}^2 \lesssim C_0^2\epsilon^2$ implies $\|z^{\mu+\nu}\|_{L^2(0,t)}^2 \lesssim \epsilon^2 + C_0\epsilon^2$ for all $(\mu, \nu) \in M$. This means that we can take $C_0 \sim 1$. This completes the proof of Proposition 6.2. □

Proof of the asymptotics (6-9). We write (6-12) in the form $\mathrm{i}\dot\eta = -\Delta\eta + V\eta + \mathbb{B}$. Then $\partial_t(e^{-\mathrm{i}\Delta t}\eta) = -\mathrm{i}e^{-\mathrm{i}\Delta t}(\eta + \mathbb{B})$ and so

$$e^{-\mathrm{i}\Delta t_2}\eta(t_2) - e^{-\mathrm{i}\Delta t_1}\eta(t_1) = -\mathrm{i}\int_{t_1}^{t_2} e^{-\mathrm{i}\Delta t}(V\eta(t) + \mathbb{B}(t))\,dt \quad \text{for } t_1 < t_2.$$

Then, for a fixed c_2, by the Strichartz estimates,

$$\|e^{-\mathrm{i}\Delta t_2}\eta(t_2) - e^{-\mathrm{i}\Delta t_1}\eta(t_1)\|_{H^1} \le c_2\big(\|\eta\|_{L^2(\mathbb{R}_+, W^{1,6})} + \|\mathbb{B}(t)\|_{L^1([t_1,t_2],H^1)+L^2([t_1,t_2],W^{6/5})}\big).$$

Since we have

$$\mathbb{B} = \sum_{(\mu,\nu)\in M} \bar z^\mu z^\nu \bar G_{\mu\nu} + \mathbb{A},$$

and by (6-14) and (6-15), valid now in $[0, \infty)$, for a fixed C we have

$$\Big\| \sum_{(\mu,\nu)\in M} \bar z^\mu z^\nu \bar G_{\mu\nu} \Big\|_{L^2(\mathbb{R}_+, W^{1,6/5})} \le C\epsilon, \quad \|\mathbb{A}\|_{L^2(\mathbb{R}_+, W^{1,6/5}) + L^1(\mathbb{R}_+, H^1)} \le C\epsilon^2,$$

so we conclude that there exists an $\eta_+ \in H^1$ with

$$\lim_{t\to+\infty} e^{-\mathrm{i}\Delta t}\eta(t) = \eta_+ \quad \text{in } H^1 \qquad \text{and} \qquad \|\eta(t) - e^{\mathrm{i}\Delta t}\eta_+\|_{H^1} \le C\epsilon \quad \text{for all } t \ge 0.$$

So we have the first limit in (6-7) and the inequality $\|\eta_+\|_{H^1} \le C\|u(0)\|_{H^1}$ in Theorem 6.1.

We prove now the existence of z_+ and the facts about it in Theorem 6.1. First of all, from (6-27),

$$\frac{1}{2}\sum_j \frac{d}{dt}|z_j|^2 = \sum_j \mathrm{Im}\Bigl[\partial_{\bar{j}}\mathcal{R}\bar{z}_j + \sum_{(\mu,\nu)\in M} \nu_j z^\mu \bar{z}^\nu \langle \eta, G_{\mu\nu}\rangle + \sum_{(\mu',\nu')\in M} \mu'_j z^{\nu'}\bar{z}^{\mu'}\langle \bar{\eta}, \bar{G}_{\mu'\nu'}\rangle\Bigr].$$

Since the right-hand side has $L^1(0,\infty)$ norm bounded by $C\epsilon^2$ for a fixed C, we conclude that the limit

$$\lim_{t\to+\infty}(|z_1(t)|, \dots, |z_n(t)|) = (\rho_{+1}, \dots, \rho_{+n}) \quad \text{exists with} \quad |\rho_+| \le C\|u(0)\|_{H^1}.$$

Since $\lim_{t\to+\infty} \boldsymbol{Z}(t) = 0$, we conclude that all but at most one of the ρ_{+j} are equal to 0. □

7. Proof of Theorem 1.4

The stability of $e^{-\mathrm{i}tE_{1z}}Q_{1z}$ is known. By [Grillakis et al. 1987, Theorem 1], the stability of $e^{-\mathrm{i}tE_{1z}}Q_{1z}$, or equivalently of $e^{-\mathrm{i}tE_{1\rho_1}}Q_{1\rho}$ for $\rho > 0$, is a consequence of the following two points:

(1) The self-adjoint operator $L_{-\rho} := H - E_{1\rho} + |Q_{1\rho}|^2$ has kernel $\ker L_{-\rho} = \{Q_{1\rho}\}$ and $L_{-\rho} > 0$ in $\{Q_{1\rho}\}^\perp$.

(2) The self-adjoint operator $L_{+\rho} = H - E_{1\rho} + 3|Q_{1\rho}|^2$ is strictly positive: $L_{+\rho} > 0$.

If $|Q_{1\rho}(x)| > 0$ for all x, then (2) is an immediate consequence of (1). The fact that $\ker L_{-\rho} = \{Q_{1\rho}\}$ follows by the facts that $Q_{1\rho} \in \ker L_{-\rho}$ and that, for $|\rho| < \epsilon_0$ with $\epsilon_0 > 0$ small, the number $E_{1\rho} \sim e_1$ is the smallest eigenvalue of $H + |Q_{1\rho}|^2$, since e_1 is the smallest eigenvalue of H.

We recall that [Tsai and Yau 2002b; 2002c; 2002d; Soffer and Weinstein 2004; Gang and Weinstein 2008; 2011; Gustafson and Phan 2011; Nakanishi et al. 2012] give partial proofs of the instability of the second excited state, and only for $2e_2 > e_1$. We now prove the instability of the excited states.

Fix $j > 1$ and assume that Q_{jr} is orbitally stable. Then Q_{jr} is asymptotically stable, by Theorem 1.3. So, if $\|u(0) - Q_{jr}\|_{H^1} \ll 1$, then $\|u(t) - Q_{jz_j} - e^{\mathrm{i}\Delta t}\eta_+\|_{H^1} \to 0$ for $t \to \infty$ and $|z_j(t)| \to \rho$ with $\rho \ne 0$ and close to r. In this case we have

$$E(u(0)) = \lim_{t\to\infty} E(u(t)) = \lim_{t\to\infty} E(Q_{jz_j(t)} + e^{\mathrm{i}\Delta t}\eta_+),$$
$$\|u(0)\|_{L^2}^2 = \lim_{t\to\infty} \|Q_{jz_j(t)} + e^{\mathrm{i}\Delta t}\eta_+\|_{L^2}^2.$$

Since $\|e^{\mathrm{i}\Delta t}\eta_+\|_{L_t^2 L_x^6} \lesssim \|\eta_+\|_{L^2}$, there exists $t_n \to \infty$ such that $\|e^{\mathrm{i}\Delta t_n}\eta\|_{L_x^6} \to 0$. So, since $\|e^{\mathrm{i}t_n\Delta}\eta_+\|_{L^4} \to 0$, $\int V|e^{\mathrm{i}t_n\Delta}\eta_+|^2\,dx \to 0$, and the cross terms in (3-3) disappear, we have

$$E(u(0)) = \lim_{n\to\infty} E(Q_{jz_j(t_n)} + e^{\mathrm{i}\Delta t_n}\eta_+) = E(Q_{j\rho}) + \|\nabla\eta_+\|_{L^2}^2,$$
$$\|u(0)\|_{L^2}^2 = \lim_{n\to\infty} \|Q_{jz_j(t_n)} + e^{\mathrm{i}\Delta t_n}\eta_+\|_{L^2}^2 = \|Q_{j\rho}\|_{L^2}^2 + \|\eta_+\|_{L^2}^2.$$

We claim that for $j \ge 2$ we can construct a curve on H^1 with the following property:

Lemma 7.1. *For sufficiently small* δ, *there exists a map* $[0,\delta)\to H^1$, $\varepsilon\mapsto\Psi(\varepsilon)$ *such that*:

- $\Psi(0)=Q_{jr}$;
- $\|\Psi(\varepsilon)\|_{L^2}^2=\|Q_{jr}\|_{L^2}^2$;
- $E(\Psi(\varepsilon))<E(Q_{jr})$ *if* $\varepsilon>0$.

Before proving the lemma, we show that the assumption that Q_{jr} is asymptotically stable and the existence of Ψ lead to a contradiction.

Proof of instability. Since $\|Q_{jr}\|_{L^2}^2=r^2+O(r^6)$ by Proposition 1.1, $\|Q_{jr}\|_{L^2}^2$ is strictly increasing in r for r small. By Proposition 1.1, we have $E'(Q_{jr})=(e_j+O(r^2))Q'(Q_{jr})$. This implies that $E(Q_{jr})$ is a strictly decreasing function of r. Setting $u(0)=\Psi(\varepsilon)$, we have

$$\|Q_{jr}\|_{L^2}^2=\|\Psi(\varepsilon)\|_{L^2}^2=\|Q_{j\rho}\|_{L^2}^2+\|\eta_+\|_{L^2}^2.$$

Therefore we have $\|Q_{jr}\|_{L^2}^2\geq\|Q_{j\rho}\|_{L^2}^2$. This implies $r\geq\rho$ and so $E(Q_{j\rho})\geq E(Q_{jr})$. But, looking at the energy, we get the following contradiction, which ends the proof of Theorem 1.4:

$$E(Q_{jr})>E(\Psi(\varepsilon))=E(Q_{j\rho})+\|\nabla\eta_+\|_{L^2}^2\geq E(Q_{j\rho})\geq E(Q_{jr}).\qquad\square$$

We now construct the curve Ψ.

Proof of Lemma 7.1. We set $\Psi(\varepsilon)=\beta(\varepsilon)Q_{j,r}+\varepsilon\phi_1$ and choose $\beta(\varepsilon)$ to make $\|\Psi(\varepsilon)\|_{L^2}^2=\|Q_{jr}\|_{L^2}^2$:

$$\|Q_{jr}\|_{L^2}^2\beta^2+2\varepsilon\langle Q_{jr},\phi_1\rangle\beta+\varepsilon^2-\|Q_{jr}\|_{L^2}^2=0.$$

So, we have

$$\begin{aligned}
\beta(\varepsilon)&=\frac{-\langle Q_{jr},\phi_1\rangle\varepsilon+\sqrt{\langle Q_{jr},\phi_1\rangle^2\varepsilon^2-\|Q_{jr}\|_{L^2}^2(\varepsilon^2-\|Q_{jr}\|_{L^2}^2)}}{\|Q_{jr}\|_{L^2}^2}=\sqrt{1-g_1(r)\varepsilon^2}+g_2(r)\varepsilon,\\
g_1(r)&:=\frac{1}{\|Q_{jr}\|_{L^2}^4}(\|Q_{jr}\|_{L^2}^2-\langle Q_{jr},\phi_1\rangle^2)=\frac{1}{\|Q_{jr}\|_{L^2}^4}(\|Q_{jr}\|_{L^2}^2-\langle q_{jr},\phi_1\rangle^2),\\
g_2(r)&:=-\frac{\langle Q_{jr},\phi_1\rangle}{\|Q_{jr}\|_{L^2}^2}=-\frac{\langle q_{jr},\phi_1\rangle}{\|Q_{jr}\|_{L^2}^2}.
\end{aligned}$$

We now show $E(\Psi(\varepsilon))<E(Q_{j,r})$ for $\varepsilon>0$. It suffices to show $S_{E_{jr}}(\Psi(\varepsilon))<S_{E_{jr}}(Q_{jr})$, where

$$S_{E_{jr}}(u)=E(u)-E_{jr}\|u\|_{L^2}^2.$$

Notice that we have $S'_{E_{jr}}(Q_{jr})=0$. Therefore, setting $\gamma(\varepsilon)=\beta(\varepsilon)-1$, we have

$$\begin{aligned}
S_{E_{jr}}(\Psi(\varepsilon))&=S_{E_{jr}}(Q_{jr}+\gamma(\varepsilon)Q_{jr}+\varepsilon\phi_1)\\
&=S_{E_{jr}}(Q_{jr})+\tfrac12\big\langle S''_{E_{jr}}(Q_{jr})(\gamma(\varepsilon)Q_{jr}+\varepsilon\phi_1),\gamma(\varepsilon)Q_{jr}+\varepsilon\phi_1\big\rangle+o(\|\gamma(\varepsilon)Q_{jr}+\varepsilon\phi_1\|_{H^1}^2).
\end{aligned}$$

If $g_2(r)=0$, we have $\gamma(\varepsilon)=O(\varepsilon^2r^{-2})$ and we conclude

$$\begin{aligned}
S_{E_{jr}}(\Psi(\varepsilon))&=S_{E_{jr}}(Q_{jr})+\varepsilon^2\langle S_{E_{jr}}(Q_{jr})\phi_1,\phi_1\rangle+o(\varepsilon^2)\\
&=S_{E_{jr}}(Q_{jr})+\varepsilon^2(e_1-e_j)+O(\varepsilon^2r)+o(\varepsilon^2)<S_{E_{jr}}(Q_{jr}).
\end{aligned}$$

If $g_2(r) \neq 0$, we have $\gamma(\varepsilon) = O(r\varepsilon)$ and

$$S_{E_{jr}}(\Psi(\varepsilon)) = S_{E_{jr}}(Q_{jr}) + \varepsilon^2(e_1 - e_j) + O(r\varepsilon^2) < S_{E_{jr}}(Q_{jr}).$$

Therefore Lemma 7.1 is proved. This also completes the proof of Theorem 1.4. $\square$

Appendix A: A generalization of Proposition 1.1

For reference purposes, we generalize (1-1) as

$$\mathrm{i}u_t = -\Delta u + V(x)u + \beta(|u|^2)u, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^3, \tag{A-1}$$

and assume that $\beta(0) = 0$, $\beta \in C^\infty(\mathbb{R}, \mathbb{R})$ and, further, there exists $p \in (1,5)$ such that, for every $k \geq 0$, there is a fixed C_k with

$$\left| \frac{d^k}{dv^k} \beta(v^2) \right| \leq C_k |v|^{p-k-1} \quad \text{if } |v| \geq 1.$$

Proposition A.1. *Fix $j \in \{1, \dots, n\}$. Then there is $a_0 > 0$ such that, for all $z_j \in B_{\mathbb{C}}(0, a_0)$, there is a unique $Q_{jz_j} \in \mathcal{S}(\mathbb{R}^3, \mathbb{C}) := \bigcap_{t \geq 0} \Sigma_t(\mathbb{R}^3, \mathbb{C})$ such that*

$$(-\Delta + V)Q_{jz_j} + \beta(|Q_{jz_j}|^2)Q_{jz_j} = E_{jz_j}Q_{jz_j}, \quad Q_{jz_j} = z_j\phi_j + q_{jz_j}, \quad \langle q_{jz_j}, \bar{\phi}_j \rangle = 0, \tag{A-2}$$

and such that we have, for any $r \in \mathbb{N}$:

(1) *$(q_{jz_j}, E_{jz_j}) \in C^\infty(B_{\mathbb{C}}(a_0), \Sigma_r \times \mathbb{R})$; we have $q_{jz_j} = z_j \hat{q}_j(|z_j|^2)$ with $\hat{q}_j(t^2) = t^2 \tilde{q}_j(t^2)$, where $\tilde{q}_j(t) \in C^\infty((-a_0{}^2, a_0{}^2), \Sigma_r(\mathbb{R}^3, \mathbb{R}))$, and $E_{jz_j} = E_j(|z_j|^2)$ with $E_j(t) \in C^\infty((-a_0{}^2, a_0{}^2), \mathbb{R})$.*

(2) *There exists $C > 0$ such that $\|q_{jz_j}\|_{\Sigma_r} \leq C|z_j|^3$ and $|E_{jz_j} - e_j| < C|z_j|^2$.*

The rest of this section is devoted to the proof of Proposition A.1.

The first step is the following lemma, which follows by a direct computation:

Lemma A.2. *Let $m \in \mathbb{N}_0$ and $k \in \{1, 2, 3\}$. Then we have*

$$\begin{aligned} [-\Delta, |x|^{2m}] &= -2m(2m+1)|x|^{2m-2} - 4m|x|^{2m-2}x \cdot \nabla, \\ [-\Delta, |x|^{2m}x_k] &= -2m(2m+3)|x|^{2m-2}x_k - 4mx_k|x|^{2m-2}x \cdot \nabla - 2|x|^{2m}\partial_{x_k}. \end{aligned} \tag{A-3}$$

Our second step is the following lemma:

Lemma A.3. *The eigenfunctions ϕ_j of $-\Delta + V$ satisfy $\phi_j \in \mathcal{S}(\mathbb{R}^3)$.*

Proof. First, $\phi_j \in L^2(\mathbb{R}^3)$, so we have $\phi_j \in H^2(\mathbb{R}^3)$ by

$$(-\Delta - e_j)\phi_j = -V\phi_j.$$

Furthermore, if we have $\phi_j \in H^{2m}(\mathbb{R}^3)$, then we have $\phi_j \in H^{2m+2}(\mathbb{R}^3)$. This implies $\phi_j \in \bigcap_{m=1}^\infty H^m$.

Next, by Lemma A.2, we have

$$(-\Delta - e_j)x_k\phi_j = -2\partial_{x_k}\phi_j - Vx_k\phi_j$$

for $k = 1, 2, 3$. Therefore, we have $x_k\phi_j \in H^2(\mathbb{R}^3)$. Again, by Lemma A.2, we have

$$(-\Delta - e_j)|x|^2\phi_j = -6\phi_j - 4x\cdot\nabla\phi_j - Vx_k\phi_j.$$

So, by $x\cdot\nabla\phi_j = \nabla(x\phi_j) - 3\phi_j \in L^2(\mathbb{R}^3)$, we have $|x|^2\phi_j \in H^2$.

Now, suppose $|x|^{2m}\phi_j \in H^2(\mathbb{R}^3)$. By Lemma A.2, we have

$$(-\Delta - e_j)|x|^{2m}x_k\phi_j = -2m(2m+3)|x|^{2m-2}x_k\phi_j - 4mx_k|x|^{2m-2}x\cdot\nabla\phi_j - 2|x|^{2m}\partial_{x_k}\phi_j - V|x|^{2m}x_k\phi_j.$$

Since

$$|x|^{2m}\partial_{x_k}\phi_j = \partial_{x_k}(|x|^{2m}\phi_j) - 4m|x|^{2m-2}x_k\phi_j \in L^2(\mathbb{R}^3),$$

we have $|x|^{2m}x_k\phi_j \in H^2(\mathbb{R}^3)$. Finally, since

$$(-\Delta - e_j)|x|^{2m+2}\phi_j = -2(m+1)(2m+3)|x|^{2m}\phi_j - 4(m+1)|x|^{2m}x\cdot\nabla\phi_j - V|x|^{2m+2}\phi_j$$

and $|x|^{2m}x\cdot\nabla\phi_j = \nabla\cdot(|x|^{2m}x\phi_j) - (4m+3)|x|^{2m}\phi_j \in L^2(\mathbb{R}^3)$, we have $|x|^{2m+2}\phi_j \in H^2(\mathbb{R}^2)$. By induction, we have $\phi_j \in \Sigma_{2m}$ for any $m \geq 1$. ☐

The next step is the following lemma:

Lemma A.4. *Fix $j \in \{1, \dots, n\}$ and $r \in \mathbb{N}$ with $r \geq 2$. Then there exists $\delta_r > 0$ such that, for all $z_j \in B_{\mathbb{C}}(0, \delta_r)$, there is a unique $Q_{jz_j} \in \Sigma_r(\mathbb{R}^3, \mathbb{C})$ satisfying (1-3) and Proposition 1.1(1)–(2).*

Proof. In this proof we write $g(u) := \beta(|u|^2)u$. Notice that it suffices to show the claim of Lemma A.4 for $z_j \in \mathbb{R}$ with real-valued Q_{j,z_j}. Indeed, if we define

$$Q_{jz_j} = e^{i\theta}Q_{j\rho} \quad \text{and} \quad E_{jz_j} = E_{j\rho} \tag{A-4}$$

for $z = e^{i\theta}\rho$, then Q_{jz} and E_{jz} satisfy (1-3) if $Q_{j\rho}$ and $E_{j\rho}$ satisfy (1-3). Further, if $B_{\mathbb{R}}(0, \delta) \to \Sigma_r \times \mathbb{R}$, $z \mapsto (Q_{jz}, E_{jz})$ is C^∞, then, by (A-4), we have $B_{\mathbb{C}}(0, \delta)\Sigma_r \times \mathbb{R}$, $z \mapsto (Q_{j,z}, E_{j,z})$ is C^∞.

Fix $j \in \{0, 1, \dots, n\}$. For simplicity we set $z_j = z$, $e_j = e$ and $\phi_j = \phi$. Set

$$Q_{j,z} = z(\phi + |z|^2\psi(z)) \quad \text{and} \quad E_{j,z} = e + |z|^2 f(z).$$

We solve (1-3) under the above ansatz. Substituting the ansatz into (1-3), we have

$$H\psi + z^{-3}g(z(\phi + z^2\psi)) = e\psi + f\phi + z^2 f\psi. \tag{A-5}$$

Set $Pu = u - \langle u, \phi\rangle\phi$. Then, we have

$$H\psi + z^{-3}Pg(z(\phi + z^2\psi)) = e\psi + z^2 f\psi, \quad \langle z^{-3}g(z(\phi + z^2\psi)), \phi\rangle = f.$$

Therefore, it suffices to solve

$$(H - e)\psi = -z^{-3}Pg(z(\phi + z^2\psi)) + z^{-1}\langle g(z(\phi + z^2\psi)), \phi\rangle\psi. \tag{A-6}$$

Now, set $\tilde{\phi}(z) := \phi + z^2\psi(z)$. Then,

$$g(z\tilde{\phi}) = \beta(z^2\tilde{\phi})z\tilde{\phi} = z^3\int_0^1 \beta'(sz^2\tilde{\phi}^2)\,ds\,\tilde{\phi}^3.$$

So, (A-6) can be rewritten as

$$(H-e)\psi = -P\left(\int_0^1 \beta'(sz^2\tilde{\phi}^2)\,ds\,\tilde{\phi}^3\right) + \langle \beta(z^2\tilde{\phi}^2)\tilde{\phi},\phi\rangle\psi. \tag{A-7}$$

To show that $z \mapsto \psi(z) \in \Sigma_r$ exists and is C^∞, we use the inverse function theorem. Set

$$\Phi(z,\psi) := -(H-e)^{-1}P\left(\int_0^1 \beta'(sz^2\tilde{\phi}^2)\,ds\,\tilde{\phi}^3\right) + \langle \beta(z^2\tilde{\phi}^2)\tilde{\phi},\phi\rangle (H-e)^{-1}\psi$$

and

$$F(z,\psi) := \psi - \Phi(z,\psi).$$

Then, $F : \mathbb{R}\times P\Sigma_r \to P\Sigma_r$ is C^∞. Next, since

$$F(0,\psi) = \psi + \beta'(0)(H-e)^{-1}P\phi^3,$$

we have

$$F(0, -\beta'(0)(H-e)^{-1}P\phi^3) = 0.$$

We now compute $F_\psi(z,\psi)$:

$$\begin{aligned}\Phi_\psi(z,\psi)h = -2z^4(H-e)^{-1}P\left(\int_0^1 \beta''(sz^2\tilde{\phi}^2)s\,ds\,\tilde{\phi}^4h\right) - 3z^2(H-e)^{-1}P\left(\int_0^1 \beta'(sz^2\tilde{\phi}^2)\,ds\,\tilde{\phi}^2h\right)\\ +2z^4\langle \beta'(z^2\tilde{\phi}^2)\tilde{\phi}^2h,\phi\rangle(H-e)\psi + z^2\langle\beta(z^2\tilde{\phi}^2)h,\phi\rangle(H-e)\psi + \langle\beta(z^2\tilde{\phi}^2)\tilde{\phi},\phi\rangle(H-e)h.\end{aligned}$$

So, we have

$$F_\psi(0,\psi)h = h.$$

Therefore, by the inverse function theorem we have the conclusion of the lemma. □

The final step is to show that the $\delta_r > 0$ can be chosen independent of r.

Lemma A.5. *Consider the Q_{jz_j} in Lemma A.4. Then there is a $\delta > 0$ such that $Q_{jz_j} \in \mathcal{S}(\mathbb{R}^3)$ for $|z_j| < \delta$.*

Proof. We use a bootstrap argument similar to the proof of Lemma A.3. We can consider the Q_{jz} given in Lemma A.4 with $r = 4$. It is enough to consider $z = \rho \in (0,\delta)$ with $\delta < \delta_4$. For $\delta > 0$ sufficiently small, we also have $E_{j\rho} < \frac{1}{2}e_j < 0$. By (A-2) we have

$$(-\Delta - E_{j\rho})Q_{j\rho} = -VQ_{j\rho} - \int_0^1 \beta'(sQ_{j\rho}^2)\,ds\,Q_{j\rho}^3. \tag{A-8}$$

We proceed as in Lemma A.3. Since the commutator term and $-VQ_{j\rho}$ are the same as in Lemma A.3, we conclude that Lemma A.5 is a consequence of the following two simple facts for $m \geq 2$:

(i) If $Q_{j\rho} \in H^m$, then $\beta(Q_{j\rho}^2)Q_{j\rho} = \int_0^1 \beta'(sQ_{j\rho}^2)\,ds\,Q_{j\rho}^3 \in H^m$.

(ii) If $|x|^{2m}Q_{j\rho} \in L^2(\mathbb{R}^3)$, then $|x|^{2m+2}\int_0^1 \beta'(sQ_{j\rho}^2)\,ds\,Q_{j\rho}^3 \in L^2$.

Fact (i) follows from the fact that $H^m(\mathbb{R}^3)$ is a ring for $m \geq 2$. We now look at (ii). Since $Q_{j\rho}$ is a continuous function with $Q_{j\rho}(x) \to 0$ as $|x| \to \infty$, the range of $Q_{j\rho}$ (i.e., $\{Q_{j\rho}(x) \in \mathbb{R} : x \in \mathbb{R}^3\}$) is relatively compact. So, since $t \to \int_0^1 \beta'(st^2)\,ds$ is a continuous function from $\mathbb{R} \to \mathbb{R}$, the range of

$\int_0^1 \beta'(sQ_{j\rho}^2)\,ds$ is relatively compact. Therefore, we have $\int_0^1 \beta'(sQ_{j\rho}^2)\,ds \in L^\infty$. On the other hand, by $Q_{j\rho} \in \Sigma_4$ we have $|x|Q_{j\rho} \in \Sigma_3 \hookrightarrow L^\infty$. Therefore, we have

$$|x|^{2m+2}\int_0^1 \beta'(sQ_{j\rho}^2)\,ds\,Q_{j\rho}^3 = \int_0^1 \beta'(sQ_{j\rho}^2)\,ds\,(|x|Q_{j\rho})^2|x|^{2m}Q_{j\rho} \in L^2(\mathbb{R}^3).$$

This proves (ii) and completes the proof of Lemma A.5. □

Finally, Proposition A.1 is a consequence of Lemmas A.2–A.5.

Appendix B: Expansions of gauge invariant functions

We prove here (3-10) and (3-12), which are direct consequences of Lemmas B.3 and B.4.

Lemma B.1. *Let $a(z) \in C^\infty(B_{\mathbb{C}}(0,\delta),\mathbb{R})$ and $a(e^{\mathrm{i}\theta}z) = a(z)$ for any $\theta \in \mathbb{R}$. Then there exists α in $\mathbb{C}^\infty([0,\delta^2);\mathbb{R})$ such that $\alpha(|z|^2) = a(z)$.*

Proof. For $z = re^{\mathrm{i}\theta}$ we have $a(z) = a(r+\mathrm{i}0)$. Since $x \mapsto a(x+\mathrm{i}0)$ is even and smooth, we have $a(x+\mathrm{i}0) = \alpha(x^2)$ with $\alpha(x)$ smooth; see [Whitney 1943]. So $a(z) = \alpha(|z|^2)$. □

Lemma B.2. *Let $\delta > 0$. Suppose $a \in C^\infty(B_{\mathbb{C}^n}(0,\delta);\mathbb{R})$ satisfies $a(e^{\mathrm{i}\theta}z_1,\dots,e^{\mathrm{i}\theta}z_n) = a(z_1,\dots,z_n)$ for all $\theta \in \mathbb{R}$ and $a(0,\dots,0) = 0$. Then, for any $M > 0$, there exists $b_{\boldsymbol{m}}$ such that*

$$a(z_1,\dots,z_n) = \sum_{j=1}^n \alpha_j(|z_j|^2) + \sum_{|m|=1} \boldsymbol{Z}^{\boldsymbol{m}} b_{\boldsymbol{m}}(z_1,\dots,z_n) + \mathscr{R}^{0,M}(z,\boldsymbol{Z}), \tag{B-1}$$

where $\alpha_j(|z_j|^2) = a(0,\dots,0,z_j,0,\dots,0)$. Furthermore, $b_{\boldsymbol{m}} \in C^\infty(B_{\mathbb{C}^n}(0,\delta);\mathbb{R})$ and satisfies

$$b_{\boldsymbol{m}}(e^{\mathrm{i}\theta}z_1,\dots,e^{\mathrm{i}\theta}z_n) = b_{\boldsymbol{m}}(z_1,\dots,z_n) \quad \text{for all } \theta \in \mathbb{R}.$$

Proof. First, we expand a as

$$a(z_1,\dots,z_n) = a(z_1,0,\dots,0) + \int_0^1 \left(\sum_{j=2}^n \partial_j a(z_1,tz_2,\dots,tz_n)z_j + \partial_{\bar{j}} a(z_1,tz_2,\dots,tz_n)\bar{z}_j\right) dt.$$

Then, by

$$a(0,z_2,\dots,z_n) = \int_0^1 \left(\sum_{j=2}^n \partial_j a(0,tz_2,\dots,tz_n)z_j + \partial_{\bar{j}} a(0,tz_2,\dots,tz_n)\bar{z}_j\right) dt,$$

we have

$$\begin{aligned} &a(z_1,\dots,z_n) \\ &= a(z_1,0,\dots,0) + a(0,z_2,\dots,z_n) + \int_0^1 \sum_{j=2}^n \big[(\partial_j a(z_1,tz_2,\dots,tz_n) - \partial_j a(0,tz_2,\dots,tz_n))z_j \\ &\qquad + (\partial_{\bar{j}} a(z_1,tz_2,\dots,tz_n) - \partial_{\bar{j}} a(0,tz_2,\dots,tz_n))\bar{z}_j\big]\,dt \end{aligned}$$

$$
\begin{aligned}
&= a(z_1, 0, \dots, 0) + a(0, z_2, \dots, z_n) \\
&\quad + \sum_{j\ge 2} \int_0^1\int_0^1 \big[(\partial_1\partial_j a(sz_1, tz_2, \dots, tz_n))z_1 z_j + (\partial_{\bar 1}\partial_j a(sz_1, tz_2, \dots, tz_n))\bar z_1 z_j \\
&\qquad\qquad + (\partial_1\partial_{\bar j} a(sz_1, tz_2, \dots, tz_n))\bar z_1 z_j + (\partial_{\bar 1}\partial_{\bar j} a(sz_1, tz_2, \dots, tz_n))\bar z_1 \bar z_j\big]\, ds\, dt.
\end{aligned}
$$

Iterating this argument first for $a(0, z_2, \dots, z_n)$ and then for $a(0, \dots, 0, z_k, \dots, z_n)$, we have

$$
\begin{aligned}
a(z_1, \dots, z_n) &= a(z_1, 0, \dots, 0) + a(0, z_2, 0, \dots, 0) + \cdots + a(0, \dots, 0, z_n) \\
&\quad + \sum_{k=1}^{n-1}\sum_{j\ge k+1} \int_0^1\int_0^1 \Big[(\partial_k\partial_j a(0, \dots, 0, sz_k, tz_{k+1}, \dots, tz_n))z_k z_j \\
&\qquad + (\partial_{\bar k}\partial_j a(0, \dots, 0, sz_k, tz_{k+1}, \dots, tz_n))\bar z_k z_j \\
&\qquad + (\partial_k\partial_{\bar j} a(0, \dots, 0, sz_k, tz_{k+1}, \dots, tz_n))z_k \bar z_j \\
&\qquad + (\partial_{\bar k}\partial_{\bar j} a(0, \dots, 0, sz_k, tz_{k+1}, \dots, tz_n))\bar z_k \bar z_j\Big]\, ds\, dt. \qquad \text{(B-2)}
\end{aligned}
$$

By Lemma B.1, there exist smooth α_j such that $\alpha_j(|z_j|^2) = a(0, \dots, 0, z_j, 0, \dots, 0)$. Furthermore, the sum of the middle two terms in the integral of (B-2) has the same form as the second term in the right-hand side of (B-1). So, it remains to handle the terms in the second and fifth lines of (B-2). Since they can be treated similarly, we focus only the second line of (B-2). Set

$$
\beta_{jk}(z_k, \dots, z_n) = \int_0^1\int_0^1 (\partial_k\partial_j a(0, \dots, 0, sz_k, tz_{k+1}, \dots, tz_n))\, ds\, dt
$$

with $j \ge k+1$. Notice that $\partial^\alpha\bar\partial^\beta a(0, \dots, 0) \ne 0$ by the gauge invariance of a is easily shown to imply $|\alpha| = |\beta|$. This in particular implies $\beta_{jk}(0, \dots, 0) = 0$. So, as in (B-2), we have

$$
\begin{aligned}
\beta_{jk}(z_k, \dots, z_n) &= \beta_{jk}(z_k, 0, \dots, 0) + \beta_{jk}(0, z_{k+1}, 0, \dots, 0) + \cdots + \beta_{jk}(0, \dots, 0, z_n) \\
&\quad + \sum_{m=k}^{n-1}\sum_{l\ge m+1} \int_0^1\int_0^1 \Big[(\partial_m\partial_l \beta_{jk}(0, \dots, 0, sz_m, tz_{m+1}, \dots, tz_n))z_m z_l \\
&\qquad + (\partial_{\bar m}\partial_l \beta_{jk}(0, \dots, 0, sz_m, tz_{m+1}, \dots, tz_n))\bar z_m z_l \\
&\qquad + (\partial_m\partial_{\bar l} \beta_{jk}(0, \dots, 0, sz_m, tz_{m+1}, \dots, tz_n))z_m \bar z_l \\
&\qquad + (\partial_{\bar m}\partial_{\bar l} \beta_{jk}(0, \dots, 0, sz_m, tz_{m+1}, \dots, tz_n))\bar z_m \bar z_l\Big]\, ds\, dt. \qquad \text{(B-3)}
\end{aligned}
$$

Since $z_l^2\beta_{jk}(0, \dots, 0, z_l, 0, \dots, 0)$ is gauge invariant by Lemma B.1, we have

$$
z_l^2\beta_{jk}(0, \dots, 0, z_l, 0, \dots, 0) = \tilde\beta_{jkl}(|z_l|^2) = \tilde\beta_{jkl}(0) + \tilde\beta'_{jkl}(0)|z_l|^2 + \gamma_{jkl}(|z_l|^2)|z_l|^4
$$

for some smooth $\tilde{\beta}_{jkl}$ and γ_{jkl}. By the smoothness of β_{jk}, we have $\tilde{\beta}_{jkl}(0) = \tilde{\beta}'_{jkl}(0) = 0$. Therefore,

$$\beta_{jk}(0, \dots, 0, z_l, 0, \dots, 0) z_k z_j = \gamma_{jkl}(|z_l|^2) z_k z_j \bar{z}_l^2 \quad \text{with} \quad k < \min\{j, l\}.$$

This can be absorbed in the second term of the right-hand side of (B-1). The same is true of the contribution of the last two lines of (B-3). The term

$$\int_0^1 \int_0^1 (\partial_m \partial_l \beta_{jk}(0, \dots, 0, s z_m, t z_{m+1}, \dots, t z_n)) z_m z_l z_j z_k \, ds \, dt \tag{B-4}$$

does not have as factors components of $\boldsymbol{Z} = (z_i \bar{z}_j)_{i \neq j}$ but it is $O(|\boldsymbol{Z}|^2)$. Treating (B-4) the way we treated the second line of (B-2), and repeating the procedure a sufficient number of times, we can express (B-4) as a sum of a summation like the second in the right-hand side of (B-1) and of a term that is $O(|\boldsymbol{Z}|^M)$ for an arbitrary M. Furthermore, notice that, since we can think of the dependence on $\boldsymbol{Z} = (z_i \bar{z}_j)_{i \neq j}$ to be polynomial, and so the remainder term $R^{0,M}(z, \boldsymbol{Z})$ in (B-1) can be thought to depend polynomially on $\boldsymbol{Z} = (z_i \bar{z}_j)_{i \neq j}$, it can be thought as the restriction of a function in $\boldsymbol{Z} \in L$. □

Lemma B.3. *Take $a(z_1, \dots, z_n)$ like in Lemma B.2. Then, for any $M > 0$, there exist smooth a_j and $b_{j\boldsymbol{m}}$ such that, for $\alpha_j(|z_j|^2) = a(0, \dots, 0, z_j, 0, \dots, 0)$, we have*

$$a(z_1, \dots, z_n) = \sum_{j=1}^{n} \alpha_j(|z_j|^2) + \sum_{1 \le |\boldsymbol{m}| \le M-1} \boldsymbol{Z}^{\boldsymbol{m}} b_{j\boldsymbol{m}}(|z_j|^2) + \mathcal{R}^{0,M}(z, \boldsymbol{Z}). \tag{B-5}$$

Proof. To prove (B-5), one only has to repeatedly use Lemma B.2. □

Lemma B.4. *Suppose that $a : \mathbb{C}^n \to \mathcal{S}$ is smooth from $B_{\mathbb{R}^{2n}}(0, \delta_r)$ to Σ_r for arbitrary $r \in \mathbb{R}$ and satisfies $a(e^{i\theta} z_1, \dots, e^{i\theta} z_n) = a(z_1, \dots, z_n)$, $a(0, \dots, 0) = 0$. Then, for any $M > 0$, there exist smooth a_j and $G_{j\boldsymbol{m}}$ such that, for $\alpha_j(|z_j|^2) = a(0, \dots, 0, z_j, 0, \dots, 0)$, we have*

$$a(z_1, \dots, z_n) = \sum_{j=1}^{n} \alpha_j(|z_j|^2) + \sum_{1 \le |\boldsymbol{m}| \le M-1} \boldsymbol{Z}^{\boldsymbol{m}} G_{j\boldsymbol{m}}(|z_j|^2) + \boldsymbol{S}^{0,M}(z, \boldsymbol{Z}). \tag{B-6}$$

Proof. The proof is same as the proof of Lemmas B.1–B.3 □

Acknowledgments

Cuccagna was partially funded by grants FIRB 2012 (Dinamiche Dispersive) from the Italian Government and FRA 2013 from the University of Trieste. Maeda was supported by the Japan Society for the Promotion of Science (JSPS) with the Grant-in-Aid for Young Scientists (B) 24740081.

References

[Bambusi and Cuccagna 2011] D. Bambusi and S. Cuccagna, "On dispersion of small energy solutions to the nonlinear Klein Gordon equation with a potential", *Amer. J. Math.* **133**:5 (2011), 1421–1468. MR 2843104 Zbl 0910.35107

[Beceanu 2012] M. Beceanu, "A critical center-stable manifold for Schrödinger's equation in three dimensions", *Comm. Pure Appl. Math.* **65**:4 (2012), 431–507. MR 2877342 Zbl 1234.35240

[Buslaev and Perel′man 1995] V. S. Buslaev and G. S. Perel′man, "On the stability of solitary waves for nonlinear Schrödinger equations", pp. 75–98 in *Nonlinear evolution equations*, edited by N. N. Ural′tseva, Amer. Math. Soc. Transl. Ser. 2 **164**, Amer. Math. Soc., Providence, RI, 1995. MR 96e:35157 Zbl 0841.35108

[Cuccagna 2003] S. Cuccagna, "On asymptotic stability of ground states of NLS", *Rev. Math. Phys.* **15**:8 (2003), 877–903. MR 2004k:35348 Zbl 1084.35089

[Cuccagna 2009] S. Cuccagna, "On instability of excited states of the nonlinear Schrödinger equation", *Phys. D* **238**:1 (2009), 38–54. MR 2010m:35482 Zbl 1161.35500

[Cuccagna 2011a] S. Cuccagna, "The Hamiltonian structure of the nonlinear Schrödinger equation and the asymptotic stability of its ground states", *Comm. Math. Phys.* **305**:2 (2011), 279–331. MR 2012d:37176 Zbl 1222.35183

[Cuccagna 2011b] S. Cuccagna, "On scattering of small energy solutions of non-autonomous Hamiltonian nonlinear Schrödinger equations", *J. Differential Equations* **250**:5 (2011), 2347–2371. MR 2012e:35237 Zbl 1216.35134

[Cuccagna 2012] S. Cuccagna, "On the Darboux and Birkhoff steps in the asymptotic stability of solitons", *Rend. Istit. Mat. Univ. Trieste* **44** (2012), 197–257. MR 3019561 Zbl 06151718

[Cuccagna 2014] S. Cuccagna, "On asymptotic stability of moving ground states of the nonlinear Schrödinger equation", *Trans. Amer. Math. Soc.* **366**:6 (2014), 2827–2888. MR 3180733 Zbl 1293.35289

[Cuccagna and Maeda 2014] S. Cuccagna and M. Maeda, "On weak interaction between a ground state and a non-trapping potential", *J. Differential Equations* **256**:4 (2014), 1395–1466. MR 3145762 Zbl 1285.35107

[Fukuizumi and Sacchetti 2011] R. Fukuizumi and A. Sacchetti, "Bifurcation and stability for nonlinear Schrödinger equations with double well potential in the semiclassical limit", *J. Stat. Phys.* **145**:6 (2011), 1546–1594. MR 2863720 Zbl 1252.82017

[Gang and Weinstein 2008] Z. Gang and M. I. Weinstein, "Dynamics of nonlinear Schrödinger/Gross–Pitaevskii equations: mass transfer in systems with solitons and degenerate neutral modes", *Anal. PDE* **1**:3 (2008), 267–322. MR 2010j:35499 Zbl 1175.35136

[Gang and Weinstein 2011] Z. Gang and M. I. Weinstein, "Equipartition of mass in nonlinear Schrödinger/Gross–Pitaevskii equations", *Appl. Math. Res. Express. AMRX* **2** (2011), 123–181. MR 2012m:35306 Zbl 1228.35224

[Goodman 2011] R. Goodman, "Hamiltonian Hopf bifurcations and dynamics of NLS/GP standing-wave modes", *J. Phys. A* **44**:42 (2011), Art. ID #425101. MR 2844997 Zbl 1161.35500

[Grecchi et al. 2002] V. Grecchi, A. Martinez, and A. Sacchetti, "Destruction of the beating effect for a non-linear Schrödinger equation", *Comm. Math. Phys.* **227**:1 (2002), 191–209. MR 2003c:81057 Zbl 1001.81011

[Grillakis et al. 1987] M. Grillakis, J. Shatah, and W. Strauss, "Stability theory of solitary waves in the presence of symmetry, I", *J. Funct. Anal.* **74**:1 (1987), 160–197. MR 88g:35169 Zbl 0589.68022

[Gustafson and Phan 2011] S. Gustafson and T. V. Phan, "Stable directions for degenerate excited states of nonlinear Schrödinger equations", *SIAM J. Math. Anal.* **43**:4 (2011), 1716–1758. MR 2012h:35325 Zbl 1230.35129

[Gustafson et al. 2004] S. Gustafson, K. Nakanishi, and T.-P. Tsai, "Asymptotic stability and completeness in the energy space for nonlinear Schrödinger equations with small solitary waves", *Int. Math. Res. Not.* **2004**:66 (2004), 3559–3584. MR 2005g:35268 Zbl 1072.35167

[Kirr et al. 2008] E. W. Kirr, P. G. Kevrekidis, E. Shlizerman, and M. I. Weinstein, "Symmetry-breaking bifurcation in nonlinear Schrödinger/Gross–Pitaevskii equations", *SIAM J. Math. Anal.* **40**:2 (2008), 566–604. MR 2010g:35299 Zbl 1157.35479

[Kirr et al. 2011] E. Kirr, P. G. Kevrekidis, and D. E. Pelinovsky, "Symmetry-breaking bifurcation in the nonlinear Schrödinger equation with symmetric potentials", *Comm. Math. Phys.* **308**:3 (2011), 795–844. MR 2012m:35308 Zbl 1235.34128

[Koo 2011] E. Koo, "Asymptotic stability of small solitary waves for nonlinear Schrödinger equations with electromagnetic potential in $\mathbb{R}^3$", *J. Differential Equations* **250**:8 (2011), 3473–3503. MR 2012f:35508 Zbl 1211.35252

[Marzuola and Weinstein 2010] J. L. Marzuola and M. I. Weinstein, "Long time dynamics near the symmetry breaking bifurcation for nonlinear Schrödinger/Gross–Pitaevskii equations", *Discrete Contin. Dyn. Syst.* **28**:4 (2010), 1505–1554. MR 2011f:35331 Zbl 1223.35288

[Nakanishi et al. 2012] K. Nakanishi, T. V. Phan, and T.-P. Tsai, "Small solutions of nonlinear Schrödinger equations near first excited states", *J. Funct. Anal.* **263**:3 (2012), 703–781. MR 2925917 Zbl 1244.35136

[Sacchetti 2005] A. Sacchetti, "Nonlinear double well Schrödinger equations in the semiclassical limit", *J. Stat. Phys.* **119**:5–6 (2005), 1347–1382. MR 2006d:81109 Zbl 1096.82014

[Soffer and Weinstein 1999] A. Soffer and M. I. Weinstein, "Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations", *Invent. Math.* **136**:1 (1999), 9–74. MR 2000k:37119 Zbl 0656.35122

[Soffer and Weinstein 2004] A. Soffer and M. I. Weinstein, "Selection of the ground state for nonlinear Schrödinger equations", *Rev. Math. Phys.* **16**:8 (2004), 977–1071. MR 2005g:81095 Zbl 1111.81313

[Taylor 1997] M. E. Taylor, *Partial differential equations, II: Qualitative studies of linear equations*, Applied Mathematical Sciences **116**, Springer, New York, 1997. MR 2011m:35002 Zbl 1206.35003

[Tsai and Yau 2002a] T.-P. Tsai and H.-T. Yau, "Asymptotic dynamics of nonlinear Schrödinger equations: resonance-dominated and dispersion-dominated solutions", *Comm. Pure Appl. Math.* **55**:2 (2002), 153–216. MR 2002i:35182 Zbl 1031.35137

[Tsai and Yau 2002b] T.-P. Tsai and H.-T. Yau, "Relaxation of excited states in nonlinear Schrödinger equations", *Int. Math. Res. Not.* **2002**:31 (2002), 1629–1673. MR 2004i:35292 Zbl 1011.35120

[Tsai and Yau 2002c] T.-P. Tsai and H.-T. Yau, "Classification of asymptotic profiles for nonlinear Schrödinger equations with small initial data", *Adv. Theor. Math. Phys.* **6**:1 (2002), 107–139. MR 2004m:35254 Zbl 1234.35030

[Tsai and Yau 2002d] T.-P. Tsai and H.-T. Yau, "Stable directions for excited states of nonlinear Schrödinger equations", *Comm. Partial Differential Equations* **27**:11–12 (2002), 2363–2402. MR 2004k:35359 Zbl 1021.35113

[Whitney 1943] H. Whitney, "Differentiable even functions", *Duke Math. J.* **10** (1943), 159–160. MR 4,193a Zbl 0063.08235

[Yajima 1995] K. Yajima, "The $W^{k,p}$-continuity of wave operators for Schrödinger operators", *J. Math. Soc. Japan* **47**:3 (1995), 551–581. MR 97f:47049 Zbl 0837.35039

Received 3 Sep 2013. Revised 11 Jan 2015. Accepted 14 May 2015.

SCIPIO CUCCAGNA: scuccagna@units.it
Department of Mathematics and Geosciences, University of Trieste, via Valerio 12/1, I-34127 Trieste, Italy

MASAYA MAEDA: maeda@math.s.chiba-u.ac.jp
Department of Mathematics and Informatics, Chiba University, Faculty of Science, Chiba 263 8522, Japan

ANALYSIS AND PDE
Vol. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1351

TRANSITION WAVES FOR FISHER–KPP EQUATIONS WITH GENERAL TIME-HETEROGENEOUS AND SPACE-PERIODIC COEFFICIENTS

GRÉGOIRE NADIN AND LUCA ROSSI

We study existence and nonexistence results for generalized transition wave solutions of space-time heterogeneous Fisher–KPP equations. When the coefficients of the equation are periodic in space but otherwise depend in a fairly general fashion on time, we prove that such waves exist as soon as their speed is sufficiently large in a sense. When this speed is too small, transition waves do not exist anymore; this result holds without assuming periodicity in space. These necessary and sufficient conditions are proved to be optimal when the coefficients are periodic both in space and time. Our method is quite robust and extends to general nonperiodic space-time heterogeneous coefficients, showing that transition wave solutions of the nonlinear equation exist as soon as one can construct appropriate solutions of a given linearized equation.

1. Introduction

We are concerned with transition wave solutions of the space-time heterogeneous reaction-diffusion equation

$$\partial_t u - \mathrm{Tr}(A(x,t)D^2u) + q(x,t)\cdot Du = f(x,t,u), \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}. \tag{1}$$

Here D and D^2 denote respectively the gradient and the Hessian with respect to the space variables. We assume that the terms in the equation are periodic in x, with the same period. The matrix field A is uniformly elliptic and the nonlinearity $f(x,t,\cdot)$ vanishes at 0 and 1. The steady states 0 and 1 are respectively unstable and stable.

When the coefficients do not depend on (x,t), Equation (1) becomes a classical homogeneous monostable reaction-diffusion equation. The pioneering works on such equations are due to Kolmogorov, Petrovski and Piskunov [Kolmogorov et al. 1937] and Fisher [1937], when $f(u) = u(1-u)$. They investigated the existence of traveling wave solutions, that is, solutions of the form $u(x,t) = \phi(x\cdot e - ct)$, with $\phi(-\infty) = 1$, $\phi(+\infty) = 0$, $\phi > 0$. The quantity $c \in \mathbb{R}$ is the speed of the wave and $e \in S^{N-1}$ is its direction. Kolmogorov, Petrovski and Piskunov [Kolmogorov et al. 1937] proved that when $A = I_N$, $q \equiv 0$ and $f = u(1-u)$, there exists $c^* > 0$ such that (1) admits traveling waves of speed c if and

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.321186, ReaDi (Reaction-Diffusion Equations, Propagation and Modelling). Luca Rossi was partially supported by GNAMPA-INdAM.
MSC2010: 35B40, 35K10, 35K57, 35P05, 35B51.
Keywords: Fisher–KPP equation, reaction-diffusion, traveling waves, generalized transition waves, generalized principal eigenvalues.

only if $c \geq c^*$. This property was extended to more general monostable nonlinearities by Aronson and Weinberger [1978]. The properties (uniqueness, stability, attractivity, decay at infinity) of these waves have been extensively studied since then.

An increasing attention has been paid to heterogeneous reaction-diffusion since the 2000s. In particular, the existence of appropriate generalizations of traveling wave solutions has been proved for various classes of heterogeneities such as shear [Berestycki and Nirenberg 1992], time periodic [Alikakos et al. 1999], space-periodic [Berestycki and Hamel 2002; Berestycki et al. 2005; Xin 1992], space-time periodic [Nolen et al. 2005; Nadin 2009], time almost periodic [Shen 1999] and time uniquely ergodic [Shen 2011b], under several types of hypotheses on the nonlinearity. Now, the topical question is to understand whether reaction-diffusion equations with general heterogeneous coefficients admit wave-like solutions or not. A generalization of the notion of traveling waves has been given by Berestycki and Hamel [2007; 2012].

Definition 1.1 [Berestycki and Hamel 2007; 2012]. A *generalized transition wave* (in the direction $e \in S^{N-1}$) is a positive time-global solution u of (1) such that there exists a function $c \in L^\infty(\mathbb{R})$ satisfying[1]

$$\lim_{x\cdot e\to-\infty} u\left(x+e\int_0^t c(s)\,ds,\,t\right)=1, \qquad \lim_{x\cdot e\to+\infty} u\left(x+e\int_0^t c(s)\,ds,\,t\right)=0, \tag{2}$$

uniformly with respect to $t \in \mathbb{R}$. The function c is called the *speed* of the generalized transition wave u, and $\phi(x,t) := u\big(x+e\int_0^t c(s)\,ds,\,t\big)$ is the associated *profile*.

The profile of a generalized transition wave satisfies

$$\lim_{x\cdot e\to-\infty}\phi(x,t)=1, \qquad \lim_{x\cdot e\to+\infty}\phi(x,t)=0, \qquad \text{uniformly with respect to } t\in\mathbb{R}.$$

It is clear that any perturbation of c obtained by adding a function with bounded integral is still a speed of u, with a different profile. Reciprocally, if $\tilde{c}$ is another speed associated with u, then it is easy to check that $t \mapsto \int_0^t (c-\tilde{c})\,ds$ is bounded. Obviously, all the notions of waves used previously when the coefficients belong to particular classes of heterogeneities can be viewed as transition waves.

The existence of such waves has been proved for one-dimensional space heterogeneous equations with ignition-type nonlinearities (that is, $f(x,u)=0$ if $u \in [0,\theta)\cup\{1\}$ and $f(x,u)>0$ if $u\in(\theta,1)$) in parallel ways by Nolen and Ryzhik [2009] and Mellet, Roquejoffre and Sire [Mellet et al. 2010], and their stability was proved in [Mellet et al. 2009]. For space heterogeneous monostable nonlinearities, when $f(x,u)>0$ if $u\in(0,1)$ and $f(x,0)=f(x,1)=0$, transition waves might not exist [Nolen et al. 2012] in general. This justified the introduction of the alternative notion of *critical traveling wave* in [Nadin 2014] for one-dimensional equations. Some existence results have also been obtained by Zlatos for partially periodic multidimensional equations of ignition-type [Zlatoš 2013].

When the coefficients only depend on t in a general way, the existence of transition waves was first proved by Shen for bistable nonlinearities [2006] (that is, nonlinearities vanishing at $u=0$ and $u=1$ but negative near these two equilibria) and for monostable equations with time uniquely ergodic coefficients

[1]For a given function $g=g(x,t)$, the condition $\lim_{x\cdot e\to\pm\infty} g(x,t)=l$ uniformly with respect to $t\in\mathbb{R}$ means that

$$\lim_{r\to+\infty}\ \sup_{\substack{\pm x\cdot e>r\\ t\in\mathbb{R}}} |g(x,t)-l|=0.$$

[2011b]. The case of general time heterogeneous monostable equations was investigated in [Nadin and Rossi 2012], where it was observed that the notions of least and upper mean play a crucial role in such frameworks.

Definition 1.2. The *least mean* (resp. the *upper mean*) *over* $\mathbb{R}$ of a function $g \in L^\infty(\mathbb{R})$ is given by

$$\lfloor g \rfloor := \lim_{T\to+\infty} \inf_{t\in\mathbb{R}} \frac{1}{T}\int_t^{t+T} g(s)\,ds \quad \left(\text{resp. } \lceil g \rceil := \lim_{T\to+\infty} \sup_{t\in\mathbb{R}} \frac{1}{T}\int_t^{t+T} g(s)\,ds\right).$$

As shown in Proposition 3.1 of [Nadin and Rossi 2012], the definitions of $\lfloor g \rfloor$, $\lceil g \rceil$ do not change if one replaces $\lim_{T\to+\infty}$ with $\sup_{T>0}$ and $\inf_{T>0}$ respectively in the above expressions; this shows that $\lfloor g \rfloor$, $\lceil g \rceil$ are well defined for any $g \in L^\infty(\mathbb{R})$. Notice that g admits a uniform mean $\langle g \rangle$, that is, $\langle g \rangle := \lim_{T\to+\infty} \frac{1}{T}\int_t^{t+T} g(s)\,ds$ exists uniformly with respect to $t \in \mathbb{R}$, if and only if $\lfloor g \rfloor = \lceil g \rceil = \langle g \rangle$. This is the case in particular when the coefficients are uniquely ergodic.

Note that if c and $\tilde{c}$ are two speeds associated with the same wave u, then $c - \tilde{c}$ has a bounded integral and thus $\lfloor c \rfloor = \lfloor \tilde{c} \rfloor$.

It is proved in [Nadin and Rossi 2012] that when $A \equiv I_N$, $q \equiv 0$ and f only depends on (t, u) and is concave and positive with respect to $u \in (0, 1)$, there exists a speed $c_* > 0$ such that, for all $\gamma > c_*$ and $|e| = 1$, Equation (1) admits a generalized transition wave with speed $c = c(t)$ in the direction e such that $\lfloor c \rfloor = \gamma$, while no such waves exist when $\gamma < c_*$.

When the coefficients not only depend on t in a general way but also on x periodically, some of the above results have been extended. Assuming in addition that the coefficients are uniquely ergodic and recurrent with respect to t and that $A \equiv I_N$, Shen [2011a] proved the existence of a quantity c_* such that, for all $\gamma > c_*$, there exists a generalized transition wave for monostable equations with speed c.

The case of space-periodic and time general monostable equations was first studied in [Rossi and Ryzhik 2014], under the additional assumption that the dependences in t and x are separated, in the sense that A and q only depend on x, periodically, while f only depends on (t, u). They proved both the existence of generalized transition waves of speed c such that $\lfloor c \rfloor > c_*$ and the nonexistence of such waves with $\lfloor c \rfloor < c_*$. Moreover, they provided a more general nonexistence result, without assuming that the dependence on x of A and q is periodic.

The aim of the present paper is to consider the general case of coefficients depending on both x and t. As in [Rossi and Ryzhik 2014], we assume the periodicity in x only for the existence result.

2. Hypotheses and results

2A. ***Statement of the main results.*** Throughout the paper, the terms in (1) will always be assumed to satisfy the following (classical) regularity hypotheses:

(3) A is symmetric and uniformly continuous, and there exist $0 < \underline{\alpha} \le \bar{\alpha}$ such that, for all $(x, t) \in \mathbb{R}^{N+1}$, $\underline{\alpha} I \le A(x, t) \le \bar{\alpha} I$.

(4) q is bounded and uniformly continuous on $\mathbb{R}^{N+1}$.

(5) f is a Caratheodory function on $\mathbb{R}^{N+1} \times [0,1]$, and there exists $\delta > 0$ such that $f(x,t,\cdot) \in W^{1,\infty}([0,1]) \cap C^1([0,\delta))$ uniformly in $(x,t) \in \mathbb{R}^{N+1}$.

The assumption that q is uniformly continuous is a technical hypothesis that is used in the proofs in order to pass to the limit in sequences of translations of the equation. It could be replaced by $\operatorname{div} q = 0$. We further assume that f is of monostable type, 0 being the unstable equilibrium and 1 being the stable one. Namely,

$$\forall (x,t) \in \mathbb{R}^{N+1}, \quad f(x,t,0) = 0, \tag{6}$$

$$\forall (x,t) \in \mathbb{R}^{N+1}, \quad f(x,t,1) = 0, \tag{7}$$

$$\forall u \in (0,1), \quad \inf_{(x,t)\in\mathbb{R}^{N+1}} f(x,t,u) > 0. \tag{8}$$

In order to derive the existence result, we need some additional hypotheses. The first one is the standard KPP condition,

$$\forall (x,t) \in \mathbb{R}^{N+1},\ u \in [0,1], \quad f(x,t,u) \le \mu(x,t)u, \tag{9}$$

where, here and in the sequel, μ denotes the function

$$\mu(x,t) := \partial_u f(x,t,0).$$

Conditions (8), (9) imply that $\inf \mu > 0$. The second condition is

$$\exists C > 0,\ \delta, \nu \in (0,1], \quad \forall x \in \mathbb{R}^N,\ t \in \mathbb{R},\ u \in (0,\delta), \quad f(x,t,u) \ge \mu(x,t)u - Cu^{1+\nu}. \tag{10}$$

Note that a sufficient condition for (10) to hold is $f(x,t,\cdot) \in C^{1+\nu}([0,\delta])$, uniformly with respect to x, t. The last condition for the existence result is

$$\exists l = (l_1,\dots,l_N) \in \mathbb{R}^N_+, \quad \forall t \in \mathbb{R},\ u \in (0,1), \quad A, q, f \text{ are } l\text{-periodic in } x, \tag{11}$$

where a function g is said to be l-periodic in x if it satisfies

$$\forall j \in \{1,\dots,N\}, \forall x \in \mathbb{R}^N, \quad g(x + l_j e_j) = g(x),$$

$(e_1,\dots,e_N)$ being the canonical basis of $\mathbb{R}^N$.

When we say that a function is a solution (or subsolution or supersolution) of (1) we always mean that it is between 0 and 1. We deal with strong solutions whose derivatives ∂_t, D, D^2 belong to some $L^p(\mathbb{R}^{N+1})$, $p \in (1,\infty)$. Many of our statements and equations, such as (1), are understood to hold almost everywhere, even if we omit to specify it, and inf, sup are used in place of ess inf, ess sup.

The main results of this paper consist of sufficient and necessary conditions for the existence of generalized transition waves, expressed in terms of their speeds.

Theorem 2.1. *Under the assumptions (3)–(11), for all $e \in S^{N-1}$, there exists $c_* \in \mathbb{R}$ such that, for every $\gamma > c_*$, there is a generalized transition wave in the direction e with a speed c such that $\lfloor c \rfloor = \gamma$.*

The minimal speed c_* we construct is explicitly given by (29), (34) and (37). A natural question is to determine whether our construction gives an optimal speed or not; that is, do generalized transition waves with speed c such that $\lfloor c \rfloor < c_*$ exist? One naturally starts by checking if our c_* coincides with the optimal speed known to exist in some particular cases, such as space-time periodic or space independent. In Section 2C we show that this is the case. The answer in the general, non-space-periodic, case is only partial. It is contained in the next theorem, where, however, we can relax the monostability hypotheses (8)–(9) by

$$\Big\lfloor \inf_{x \in \mathbb{R}^N} \mu(x, \cdot) \Big\rfloor > 0, \tag{12}$$

and we can drop (7), (10) as well as (11). We actually need an extra regularity assumption on A:

$$A \text{ is uniformly Hölder-continuous in } x, \text{ uniformly with respect to } t. \tag{13}$$

This ensures the validity of some a priori Lipschitz estimates quoted from [Porretta and Priola 2013] that will be needed in the sequel. It is not clear to us if such estimates hold without (13).

Theorem 2.2. *Under the assumptions (3)–(6), (12)–(13), for all $e \in S^{N-1}$, there exists $c^* \in \mathbb{R}$ such that if c is the speed of a generalized transition wave in the direction e then $\lfloor c \rfloor \geq c^*$.*

We point out that no spatial-periodicity condition is assumed in the previous statement. In order to prove Theorem 2.2 we derive a characterization of the least mean — Proposition 4.4 below — that we believe to be of independent interest. The definition of c^* is given in Section 4. Of course, $c^* \leq c_*$ if the hypotheses of both Theorems 2.1 and 2.2 are fulfilled. We do not know if, in general, $c_* = c^*$, that is, if the speed c_* is minimal, in the sense that there does not exist any wave with a speed having a smaller least mean. When the coefficients are periodic in space and time or only depend on time, we could identify the speed c_* more explicitly (see Section 2C below). Indeed, we recover in these frameworks some characterizations of the speeds identified in earlier papers [Nadin 2009; Nadin and Rossi 2012; Rossi and Ryzhik 2014], which were proved to be minimal. In the general framework, we leave this question open.

Finally, we leave as an open problem the case $\lfloor c \rfloor = c_*$, for which we believe that generalized transition waves still exist.

2B. *Optimality of the monostability assumption.* The assumption (8) implies that 0 and 1 are respectively unstable and stable. Let us discuss the meaning and the optimality of this hypothesis, which might seem strong. Actually, as we do not make any additional assumption on the coefficients, we can consider much more general asymptotic states $p_- = p_-(x,t) < p_+ = p_+(x,t)$ in place of 0 and 1 and try to construct generalized transition waves v connecting p_- to p_+. Indeed, if $p_\pm$ are solutions to (1), with $p_+ - p_-$ bounded and having positive infimum, then the change of variables

$$u(x,t) := \frac{v(x,t) - p_-(x,t)}{p_+(x,t) - p_-(x,t)}$$

leads to an equation of the same form, with reaction term

$$\tilde{f}(x,t,u) := \frac{f(x,t,up_+ + (1-u)p_-) - uf(x,t,p_+) - (1-u)f(x,t,p_-)}{p_+ - p_-}.$$

The new equation admits the steady states 0 and 1. Moreover, assuming that $u \mapsto f(x,t,u)$ is strictly concave, then $\tilde{f}$ satisfies conditions (8), (9), the latter following from the inequality

$$\forall u \in (0,1), \quad u(p_+ - p_-)\partial_u f(x,t,p_-) \geq f(x,t,up_+ + (1-u)p_-) - f(x,t,p_-).$$

This shows that, somehow, the concavity hypothesis of the nonlinearity with respect to u is stronger, up to some change of variables, than the positivity hypothesis of the nonlinearity.

Let us illustrate the above procedure with an explicit example where $p_- \equiv 0$. Consider the equation

$$\partial_t v = \Delta v + \mu(x,t)v - v^2, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}, \tag{14}$$

with μ periodic in x, bounded and such that $\inf \mu > 0$. The later condition implies that the solution 0 is linearly unstable (actually, it can be relaxed by (12); see the discussion below). Then one can check that there is a time-global solution $p = p(x,t)$ which is bounded, has a positive infimum and is periodic in x. Let $u := v/p$. This function satisfies

$$\partial_t u = \Delta u + 2\frac{\nabla p}{p} \cdot \nabla u + p(x,t)u(1-u),$$

which is an equation of the form (1) for which (9)–(11) hold, at least if, for instance, μ is uniformly Hölder-continuous, since then ∇p is bounded by Schauder's parabolic estimates, and $\inf p > 0$.

Following this example, one can wonder whether (8) is an optimal condition (up to some change of variables) for the existence of transition waves. It is well-known that other classes of nonlinearities, such as bistable or ignition ones, could still give rise to transition waves (see for instance [Berestycki and Hamel 2002]). Thus, this question only makes sense if one reduces to the class of nonlinearities which are monostable, in a sense. Let us assume that f satisfies (6), (7) and that 0 is linearly unstable, in the weak sense that (12) holds. Then, using the properties of the least mean derived in [Nadin and Rossi 2012], one can construct arbitrarily small subsolutions $\underline{u} = \underline{u}(t)$ and thus, as 1 is a positive solution, there exists a minimal solution p of (1) in the class of bounded solutions with positive infimum. One could then check that our proof still works and gives rise to generalized transition waves connecting 0 to p. Indeed, condition (8) only ensures that $p \equiv 1$. As a conclusion, the positivity hypothesis (8) is not optimal: one could replace it by (12) but then the generalized transition waves we construct connect 0 to the minimal time-global solution, which might not be 1.

Since for the existence of positive solutions it is sufficient to require (12) rather than $\inf \mu > 0$, one may argue that, in order to guarantee that 1 is the minimal time-global solution with positive infimum, hypothesis (8) could be relaxed by

$$\forall u \in (0,1), \quad \left\lfloor \min_{x \in \mathbb{R}^N} f(x,\cdot,u) \right\rfloor > 0. \tag{15}$$

This is not true, as shown by the following example. Let $p \in C^1(\mathbb{R})$ be a strictly decreasing function such that $p(\pm\infty) \in (0,1)$. Let f satisfy $f(t,p(t)) = p'(t)$. It is clear that f can be extended in such a way that (15) holds; however p is a time-global solution of $\partial_t u = f(t,u)$ with positive infimum which is smaller than 1.

Finally, if 0 is linearly stable, in the sense that

$$\left\lceil \sup_{x\in\mathbb{R}^N} \mu(x,\cdot)\right\rceil < 0 \tag{16}$$

holds, and (9) is satisfied, then there do not exist generalized transition waves at all, and, more generally, solutions to the Cauchy problem with bounded initial data converge uniformly to 0 as $t\to\infty$. Indeed, as an easy application of the property of the least (and upper) mean (39), one can construct a supersolution $\bar{u} = e^{\sigma(t)-\varepsilon t}$, for some $\sigma \in W^{1,\infty}(\mathbb{R})$ and $\varepsilon > 0$. The convergence to 0 of bounded solutions then follows from the comparison principle.

2C. ***Description of the method and application to particular cases.*** The starting point of the construction of generalized transition waves consists of finding an explicit expression for the speed. This is not a trivial task in the case of mixed space-time dependence considered in this paper. We achieve it by a heuristic argument that we now illustrate.

Suppose that u is a generalized transition wave in a direction $e \in S^{N-1}$. Its tail at large $x\cdot e$ is close from being a solution of the linearized equation around 0:

$$\partial_t u - \mathrm{Tr}(A(x,t)D^2u) + q(x,t)\cdot Du = \mu(x,t)u. \tag{17}$$

It is natural to expect the tail of u to decay exponentially. Thus, since the equation is spatially periodic, we look for (the tail of) u under the form

$$u(x,t) = e^{-\lambda x\cdot e}\eta_\lambda(x,t), \quad \text{with } \eta_\lambda \text{ positive and } l\text{-periodic in } x. \tag{18}$$

Rewriting this expression as

$$u(x,t) = \exp\Big(-\lambda\Big(x\cdot e - \frac{1}{\lambda}\ln \eta_\lambda(x,t)\Big)\Big)$$

shows that the speed of u, namely, a function c for which (2) holds, should satisfy

$$\left|\int_0^t c(s)\,ds - \frac{1}{\lambda}\ln\eta_\lambda(x,t)\right| \le C,$$

for some C independent of $(x,t)\in\mathbb{R}^N\times\mathbb{R}$. Clearly, this can hold true only if the ratio between maximum and minimum of $\eta_\lambda(\,\cdot\,,t)$ is bounded uniformly on t. This property follows from a Harnack-type inequality, Lemma 3.1 below, which is the keystone of our proof and actually the only step where the periodicity in x really plays a role. It would be then natural to define $c(t) := \frac{1}{\lambda}\frac{d}{dt}\ln\|\eta_\lambda(\,\cdot\,,t)\|_{L^\infty(\mathbb{R})}$. The problem is that we do not know if this function is bounded, since it is not clear whether $\partial_t\eta_\lambda \in L^\infty(\mathbb{R}^{N+1})$ or not. We overcome this difficulty by showing that there exists a Lipschitz continuous function S_λ such that

$$\exists\beta>0,\quad \forall t\in\mathbb{R},\quad \left|S_\lambda(t) - \frac{1}{\lambda}\ln\|\eta_\lambda(\,\cdot\,,t)\|_{L^\infty(\mathbb{R}^N)}\right| \le \beta. \tag{19}$$

We deduce that the function c defined (almost everywhere) by $c := S_\lambda'$ is bounded and it is an admissible speed for the wave u. The method described above provides, for any given $\lambda>0$, a wave with speed $c=c_\lambda$ for the *linearized* equation which decays with exponential rate λ. It is known — for instance in the case of constant coefficients — that only decaying rates which are "not too fast" are admissible for waves of the nonlinear reaction-diffusion equation. In Section 3C, we identify a threshold rate λ_*. In the following section we construct generalized transition waves for any $\lambda < \lambda_*$, recovering with the least mean of their speeds the whole interval $(\lfloor c_{\lambda_*}\rfloor, +\infty)$. We do not know if the critical speed $c_* := \lfloor c_{\lambda_*}\rfloor$ is optimal, nor if an optimal speed does exist. However, we show below that this is the case if one applies the above procedure to some particular classes of heterogeneities already investigated in the literature.

In the case where the coefficients are periodic in time too, the class of admissible speeds has been characterized in [Nadin 2009] (see also [Berestycki et al. 2008]). Following the method described above, we see that an entire solution of (17) in the form (18) is given by $\eta_\lambda(x,t) = e^{k(\lambda)t}\varphi_\lambda(x,t)$, where $(k(\lambda), \varphi_\lambda)$ are the principal eigenelements of the problem[2]

$$\begin{cases} \partial_t\varphi_\lambda - \mathrm{Tr}(AD^2\varphi_\lambda) + (q+2\lambda Ae)D\varphi_\lambda - (\mu+\lambda^2 eAe + \lambda q\cdot e)\varphi_\lambda + k(\lambda)\varphi_\lambda = 0 \ \text{ in } \mathbb{R}^N\times\mathbb{R}, \\ \varphi_\lambda > 0, \\ \varphi_\lambda \text{ is periodic in } t \text{ and } x. \end{cases} \tag{20}$$

Actually, the uniqueness up to a multiplicative constant of solutions of (17) in the form (18), provided by Lemma 3.1 (proved without assuming the time-periodicity), implies that η_λ necessarily has this form. Thus, $S_\lambda(t) := (k(\lambda)/\lambda)t$ satisfies (19), whence the speed of the wave for the linearized equation with decaying rate λ is $c_\lambda \equiv k(\lambda)/\lambda$. Since the c_λ are constant (and therefore they have uniform mean), it turns out that the threshold λ_* we obtain for the decaying rates coincides with the minimum point of $\lambda \mapsto c_\lambda$ (see Remark 1 below). We eventually derive the existence of a generalized transition wave for any speed larger than $c_* := \min_{\lambda>0} k(\lambda)/\lambda$, which is exactly the sharp critical speed for pulsating traveling fronts obtained in [Nadin 2009]. To sum up, our construction of the minimal speed c_* is optimal in the space-time periodic framework. On the other hand, in the periodic framework, the speed c^* constructed in Section 4 is identical to c_* and thus Theorem 2.2 implies that there do not exist generalized transition waves with a speed c such that $\lfloor c\rfloor < \min_{\lambda>0} k(\lambda)/\lambda$. We therefore recover also the nonexistence result for pulsating traveling fronts. Only the existence of fronts with critical speed is not recovered.

In the case investigated in [Nadin and Rossi 2012], namely, when $A \equiv I_N$, $q \equiv 0$ and f does not depend on x, one can easily check that $\eta_\lambda(t) = e^{\int_0^t \mu(s)\,ds + \lambda^2 t}$. As a function S_λ we can simply take $\frac{1}{\lambda}\ln\|\eta_\lambda(\cdot,t)\|_{L^\infty(\mathbb{R}^N)} = \frac{1}{\lambda}\int_0^t \mu(s)\,ds + \lambda t$, which is Lipschitz continuous. Hence $c_\lambda(t) := \lambda + \mu(t)/\lambda$ is a speed of a wave with decaying rate λ. In this case the critical decaying rate λ_* is equal to $\sqrt{\lfloor\mu\rfloor}$ (see again Remark 1) and thus we have $c_* = 2\sqrt{\lfloor\mu\rfloor}$. This is the same speed c_* as in [Nadin and Rossi 2012], which was proved to be minimal.

Under the assumptions made in [Rossi and Ryzhik 2014], that is, A and q only depend on x (periodically) and f only depends on (t,u), the speeds c_* derived in the present paper and in [Rossi and Ryzhik 2014] coincide, and thus it is minimal, in the sense that there do not exist any generalized transition waves with a lower speed.

[2] The properties of these eigenelements, which are unique (up to a multiplicative constant in the case of φ_λ) are described in [Nadin 2009] for instance.

When $A \equiv I_N$ and q, f are periodic in x and uniquely ergodic in t, then one can prove that the same holds true for the function $\partial_t \eta_\lambda / \eta_\lambda$ by uniqueness, and thus $\alpha\lfloor c_\alpha \rfloor$ could be identified with the Lyapounov exponent $\lambda(\alpha, \xi)$ used by Shen in [2011a], where ξ is the direction of propagation. We thus recover in this framework the same speed c_* as in [Shen 2011a], which was not proved to be minimal since the nonexistence of transition waves with lower speed were not investigated. Note that this identification is not completely obvious. However, as the formalism of the present paper and [Shen 2011a] are very different, we leave these computations to the reader.

Lastly, let us consider the following example, where one could indeed construct directly the generalized transition waves:

$$\partial_t u - \partial_{xx} u - q(t)\partial_x u = \mu_0 u(1-u), \tag{21}$$

with q bounded and uniformly continuous and $\mu_0 > 0$. This equation satisfies assumptions (3)–(12). The change of variables $v(x,t) := u\big(x - \int_0^t q, t\big)$ leads to the classical homogeneous Fisher–KPP equation $\partial_t v - \partial_{xx} v = \mu_0 v(1-v)$. This equation admits traveling wave solutions of the form $v(x,t) = \phi_c(x - ct)$, with $\phi_c(-\infty) = 1$ and $\phi_c(+\infty) = 0$, for all $c \geq 2\sqrt{\mu_0}$. Hence, Equation (21) admits generalized transition waves $u(x,t) = \phi_c\big(x - ct + \int_0^t q, t\big)$ of speed $c - q(t)$ if and only if $c \geq 2\sqrt{\mu_0}$. That is, the set of least mean of admissible speeds is $[2\sqrt{\mu_0} - \lceil q \rceil, +\infty)$. Computing c_* in this case, one easily gets

$$\eta_\lambda = \eta_\lambda(t) = e^{\lambda^2 t - \lambda \int_0^t q(s)\,ds + \mu_0 t}, \quad c_\lambda(t) = \lambda - q(t) + \mu_0/\lambda \quad \text{and} \quad c_* = 2\sqrt{\mu_0} - \lceil q \rceil.$$

One could check that c^* coincides with this value too, meaning that Theorems 2.1 and 2.2 fully characterize the possible least means for admissible speeds, except for the critical one.

3. Existence result

Throughout this section, we fix $e \in S^{N-1}$ and we assume that conditions (3)–(11) hold. Actually, condition (8) could be weakened by (12), except for the arguments in the very last part of the proof in Section 3D. As already mentioned in Section 2B, these arguments could be easily adapted to the case where (8) is replaced by (12), leading to transition waves connecting 0 to the minimal solution with positive infimum.

3A. ***Solving the linearized equation.*** We focus on solutions with prescribed spatial exponential decay.

Lemma 3.1. *For all $\lambda > 0$, the equation (17) admits a time-global solution of the form (18). Moreover, η_λ is unique up to a multiplicative constant and satisfies, for all $t \in \mathbb{R}$, $T \geq 0$,*

$$\max_{x\in\mathbb{R}^N} \eta_\lambda(x, t+T) \leq \max_{x\in\mathbb{R}^N} \eta_\lambda(x,t)\, \exp\bigg(\Big(\overline{\alpha}\lambda + \sup_{\mathbb{R}^{N+1}} |q|\Big)\lambda T + \int_t^{t+T} \max_{x\in\mathbb{R}^N} \mu(x,s)\,ds\bigg), \tag{22}$$

$$\min_{x\in\mathbb{R}^N} \eta_\lambda(x, t+T) \geq C \max_{x\in\mathbb{R}^N} \eta_\lambda(x,t)\, \exp\bigg(\Big(\underline{\alpha}\lambda - \sup_{\mathbb{R}^{N+1}} |q|\Big)\lambda T + \int_t^{t+T} \min_{x\in\mathbb{R}^N} \mu(x,s)\,ds\bigg), \tag{23}$$

with $C > 0$ only depending on a constant bounding $|\lambda|$, $|l|$, $\underline{\alpha}^{-1}$, $\overline{\alpha}$, N and the L^∞ norms of μ and q.

The function $(x,t) \mapsto e^{-\lambda x\cdot e}\eta_\lambda(x,t)$ is a solution of the linearization of (1) near the unstable equilibrium. We will show in the next section that it is somehow a transition wave solution of the linearized equation, in the sense that it moves in the direction e with a certain speed. Due to hypothesis (9), we could use it as a supersolution of the nonlinear equation. Then, in Section 3C, in order to construct an appropriate subsolution, we will need to restrict to exponents λ less than some threshold λ_*. We will eventually derive the existence of transition waves in Section 3D.

As mentioned in Section 2C, Lemma 3.1 is the only point where the spatial periodicity hypothesis (11) is used. If the coefficients depend in a general way on both x and t and if one is able to construct a solution η_λ of equation (25) for which there exists $C > 0$ such that, for all $T > 0$, $(x, t) \in \mathbb{R}^{N+1}$, one has

$$\frac{1}{C}\|\eta_\lambda(\cdot\,,t)\|_{L^\infty(\mathbb{R}^N)}e^{-CT} \le \eta_\lambda(x, t+T) \le C\|\eta_\lambda(\cdot\,,t)\|_{L^\infty(\mathbb{R}^N)}e^{CT}. \tag{24}$$

Then the forthcoming other steps of the proof still apply and it is possible to construct a generalized transition wave solution of the nonlinear equation (1). We describe this extension in Section 3E below. It would be very useful to determine optimal conditions on the coefficients enabling the derivation of a global Harnack-type inequality such as (24) for the linearized equation. We leave this question as an open problem.

Proof of Lemma 3.1. The problem for η_λ is

$$\partial_t\eta_\lambda = \mathrm{Tr}(AD^2\eta_\lambda) - (q + 2\lambda Ae)\cdot D\eta_\lambda + (\mu + \lambda^2 eAe + \lambda q\cdot e)\eta_\lambda, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}. \tag{25}$$

We find a positive, l-periodic solution to (25) as the locally uniform limit of (a subsequence of) solutions η^n of the problem in $\mathbb{R}^N \times (-n, +\infty)$, with initial datum $\eta^n(-n, \cdot\,) \equiv m_n$, where m_n is a positive constant chosen in such a way that, say, $\sup_{x\in\mathbb{R}^N} \eta^n(0, x) = 1$.

Let us show that any l-periodic solution η_λ to (25) satisfies (22) and (23). For a given $t_0 \in \mathbb{R}$, the function

$$\max_{x\in\mathbb{R}^N} \eta_\lambda(x, t_0)\ \exp\biggl(\Bigl(\bar\alpha\lambda^2 + \sup_{\mathbb{R}^{N+1}} |q|\lambda\Bigr)(t-t_0) + \int_{t_0}^{t} \max_{x\in\mathbb{R}^N} \mu(x, s)\,ds\biggr)$$

is a supersolution of (25) larger than η_λ at time t_0. Since η_λ is bounded, we can apply the parabolic comparison principle and derive (22). Let $\mathcal{C}$ denote the periodicity cell $\prod_{j=1}^N [0, l_j]$. By the parabolic Harnack inequality (see, e.g., Corollary 7.42 in [Lieberman 1996]), we have that

$$\forall t \in \mathbb{R}, \quad \max_{x\in\mathcal{C}} \eta_\lambda(x, t-1) \le \tilde C \min_{x\in\mathcal{C}} \eta_\lambda(x, t), \tag{26}$$

for some $\tilde C > 0$ depending on a constant bounding $|\lambda|$, $|l|$, $\underline{\alpha}^{-1}$, $\bar\alpha$, N and the L^∞ norms of μ and q, and not on t. On the other hand, the comparison principle yields, for $T \ge 0$,

$$\min_{x\in\mathbb{R}^N} \eta_\lambda(x, t+T) \ge \min_{x\in\mathbb{R}^N} \eta_\lambda(x, t)\ \exp\biggl(\Bigl(\underline{\alpha}\lambda^2 - \sup_{\mathbb{R}^{N+1}} |q|\lambda\Bigr)T + \int_t^{t+T} \min_{x\in\mathbb{R}^N} \mu(x, s)\,ds\biggr).$$

Combining this inequality with (26) we eventually derive

$$\min_{x\in\mathbb{R}^N} \eta_\lambda(x, t+T) \ge \tilde C^{-1} \max_{x\in\mathbb{R}^N} \eta_\lambda(x, t-1)\ \exp\biggl(\Bigl(\underline{\alpha}\lambda^2 - \sup_{\mathbb{R}^{N+1}} |q|\lambda\Bigr)T + \int_t^{t+T} \min_{x\in\mathbb{R}^N} \mu(x, s)\,ds\biggr),$$

from which (23) follows by (22).

It remains to prove the uniqueness result. Assume that (17) admits two solutions η^1, η^2 that are positive and l-periodic in x. As shown before, we know that they both satisfy (22) and (23). We first claim that there exists $K > 1$ such that

$$\forall t \in \mathbb{R},\ x \in \mathbb{R}^N, \quad K^{-1}\eta^2(x,t) \le \eta^1(x,t) \le K\eta^2(x,t). \tag{27}$$

Let $h > 0$ be such that $\eta^1 \le h\eta^2$ at $t = 0$. It follows, for $t \le 0$, that $\min_{x\in\mathbb{R}^N} \eta^1(x,t) \le h \max_{x\in\mathbb{R}^N} \eta^2(x,t)$, because otherwise the parabolic strong maximum principle would imply $\eta^1 > h\eta^2$ at $t = 0$. Hence, applying (23) with $T = 0$ to both η^1 and η^2, we find a positive constant K such that

$$\forall t < 0, \quad \max_{x\in\mathbb{R}^N} \eta^1(x,t) \le K \min_{x\in\mathbb{R}^N} \eta^2(x,t).$$

This proves the second inequality in (27), for $t < 0$, whence for all $t \in \mathbb{R}$ by the maximum principle. The first inequality, with a possibly larger K, is obtained by exchanging the roles of η^1 and η^2. Now define

$$k := \limsup_{t\to-\infty} \max_{x\in\mathbb{R}^N} \frac{\eta^1(x,t)}{\eta^2(x,t)}.$$

We know from (27) that $k \in [K^{-1}, K]$. Consider a sequence $(t_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} t_n = -\infty, \quad \lim_{n\to\infty} \max_{x\in\mathbb{R}^N} \frac{\eta^1(x,t_n)}{\eta^2(x,t_n)} = k.$$

Define the sequences of functions $(\eta_n^1)_{n\in\mathbb{N}}$, $(\eta_n^2)_{n\in\mathbb{N}}$ as follows:

$$\forall i \in \{1,2\},\ n \in \mathbb{N}, \quad \eta_n^i(x,t) := \frac{\eta^i(x,t+t_n)}{\max_{y\in\mathbb{R}^N} \eta^1(y,t_n)}.$$

We deduce from (22) and (23) that the $(\eta_n^1)_{n\in\mathbb{N}}$ are uniformly bounded from above and uniformly bounded from below away from 0 in, say, $\mathbb{R}^N \times [-2,2]$. The same is true for $(\eta_n^2)_{n\in\mathbb{N}}$ by (27). Thus, by parabolic estimates and periodicity in x, the sequences $(\eta_n^i)_n$, $(\partial_t\eta_n^i)_n$, $(D\eta_n^i)_n$ and $(D^2\eta_n^i)_n$ converge, up to subsequences, in $L^p_{\mathrm{loc}}(\mathbb{R}^{N+1})$. Morrey's inequality yields that the sequences $(\eta_n^1)_n$ and $(\eta_n^2)_n$ converge locally uniformly to some functions $\tilde\eta^1$ and $\tilde\eta^2$ respectively.

Define $A_n := A(\cdot,\cdot+t_n)$, $q_n := q(\cdot,\cdot+t_n)$, $\mu_n := \mu(\cdot,\cdot+t_n)$. As A and q are uniformly continuous, $(A_n)_n$ and $(q_n)_n$ converge (up to subsequences) to some functions $\tilde A$ and $\tilde q$ in $L^\infty_{\mathrm{loc}}(\mathbb{R}^{N+1})$, whereas $(\mu_n)_n$ converges to some $\tilde\mu$ in the $L^\infty(\mathbb{R}^{N+1})$ weak-$*$ topology. Hence, taking the weak $L^p_{\mathrm{loc}}(\mathbb{R}^{N+1})$ limit as $n \to \infty$ in the equations satisfied by the $(\eta_n^i)_{n\in\mathbb{N}}$, we get

$$\partial_t\tilde\eta^i = \mathrm{Tr}(\tilde A D^2\tilde\eta^i) - (\tilde q + 2\lambda\tilde A e)D\tilde\eta^i + (\tilde\mu + \lambda^2 e\tilde A e + \lambda\tilde q\cdot e)\tilde\eta^i, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}.$$

Clearly, these equations hold almost everywhere because all the terms are measurable functions. That is, the $\tilde\eta^i$ are strong solutions. Moreover,

$$\tilde\eta^1 \le k\tilde\eta^2, \quad \max_{x\in\mathbb{R}^N} \frac{\tilde\eta^1(x,0)}{\tilde\eta^2(x,0)} = k.$$

The strong maximum principle then yields $\tilde{\eta}^1 \equiv k\tilde{\eta}^2$. As a consequence, for any $\varepsilon > 0$, we can find $n_\varepsilon \in \mathbb{N}$ such that, for $n \ge n_\varepsilon$, one has $(k-\varepsilon)\eta_n^2 < \eta_n^1 < (k+\varepsilon)\eta_n^2$ at $t = 0$. These inequalities hold for all $t \ge 0$, again by the maximum principle. Reverting to the original functions we obtain $(k-\varepsilon)\eta^2 < \eta^1 < (k+\varepsilon)\eta^2$ for $t \ge t_n$ and $n \ge n_\varepsilon$, from which, letting $n \to \infty$ and $\varepsilon \to 0^+$, we eventually infer that $\eta^1 \equiv k\eta^2$ for all $t \in \mathbb{R}$. □

In the particular case $T = 0$, the inequality (23) reads

$$\min_{x\in\mathbb{R}^N} \eta_\lambda(x,t) \ge C \max_{x\in\mathbb{R}^N} \eta_\lambda(x,t). \tag{28}$$

Notice that, in contrast with the standard parabolic Harnack inequality, the two sides are evaluated at the same time. This particular instance of (23) will be used in the sequel.

Until the end of the proof of Theorem 2.1, for $\lambda > 0$, we let η_λ stand for the (unique up to a multiplicative constant) function given by Lemma 3.1.

3B. *The speeds of the waves.*

Lemma 3.2. *There is a uniformly Lipschitz-continuous function $S_\lambda : \mathbb{R} \to \mathbb{R}$ satisfying* (19).

Proof. Properties (22)–(23) yield the existence of a constant $\beta > 0$ such that

$$\forall t \in \mathbb{R},\ T \ge 0, \quad \big|\ln \|\eta_\lambda(\cdot\,, t+T)\|_{L^\infty(\mathbb{R}^N)} - \ln \|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}\big| \le \beta(1+\lambda^2)T.$$

For all $n \in \mathbb{N}$, we define S_λ on $[n, n+1]$ as the affine function satisfying

$$S_\lambda(n) = \frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, n)\|_{L^\infty(\mathbb{R}^N)}, \qquad S_\lambda(n+1) = \frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, n+1)\|_{L^\infty(\mathbb{R}^N)}.$$

Then, for all $t \in (n, n+1)$,

$$|S_\lambda'(t)| = \left|\frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, n+1)\|_{L^\infty(\mathbb{R}^N)} - \frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, n)\|_{L^\infty(\mathbb{R}^N)}\right| \le \beta \frac{1+\lambda^2}{\lambda}.$$

Hence, S_λ is uniformly Lipschitz-continuous over $\mathbb{R}$. Moreover, if $t \in [n, n+1]$, one has

$$\begin{aligned}\left|S_\lambda(t) - \frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}\right| &\le |S_\lambda(t) - S_\lambda(n)| + \frac{1}{\lambda}\big|\ln \|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)} - \ln \|\eta_\lambda(\cdot\,, n)\|_{L^\infty(\mathbb{R}^N)}\big| \\ &\le 2\beta \frac{1+\lambda^2}{\lambda}.\end{aligned}$$

Hence, $t \mapsto S_\lambda(t) - \frac{1}{\lambda} \ln \|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}$ is uniformly bounded over $\mathbb{R}$. □

Owing to Lemma 3.2, the function c_λ, defined for (almost everywhere) $t \in \mathbb{R}$ by

$$c_\lambda(t) := S_\lambda'(t), \tag{29}$$

belongs to $L^\infty(\mathbb{R})$. We will use it as a possible speed for a transition wave to be constructed.

Let us investigate the properties of the least mean of the $(c_\lambda)_{\lambda>0}$. It follows from (19) that

$$\lfloor c_\lambda \rfloor = \frac{1}{\lambda} \lim_{T\to+\infty} \inf_{t\in\mathbb{R}} \frac{1}{T} \ln \frac{\|\eta_\lambda(\cdot\,, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}}. \tag{30}$$

Hence, by (22) and (23), we derive

$$\underline{\alpha}\lambda - \sup_{\mathbb{R}^{N+1}} |q| + \frac{1}{\lambda}\Big\lfloor \min_{x\in\mathbb{R}^N} \mu(x,\cdot)\Big\rfloor \le \lfloor c_\lambda \rfloor \le \bar{\alpha}\lambda + \sup_{\mathbb{R}^{N+1}} |q| + \frac{1}{\lambda}\Big\lfloor \max_{x\in\mathbb{R}^N} \mu(x,\cdot)\Big\rfloor. \tag{31}$$

Analogous bounds hold for the upper mean:

$$\underline{\alpha}\lambda - \sup_{\mathbb{R}^{N+1}} |q| + \frac{1}{\lambda}\Big\lceil \min_{x\in\mathbb{R}^N} \mu(x,\cdot)\Big\rceil \le \lceil c_\lambda \rceil \le \bar{\alpha}\lambda + \sup_{\mathbb{R}^{N+1}} |q| + \frac{1}{\lambda}\Big\lceil \max_{x\in\mathbb{R}^N} \mu(x,\cdot)\Big\rceil. \tag{32}$$

We have seen in Section 2C that, when the coefficients are periodic in t, one can take $S_\lambda(t) := (k(\lambda)/\lambda)t$, whence $c_\lambda \equiv k(\lambda)/\lambda$. It follows that $\lambda c_\lambda = k(\lambda)$, and we know from the arguments in the proof of Proposition 5.7 part (iii) in [Berestycki and Hamel 2002] that the function k is convex. In the general heterogeneous framework considered in the present paper, we use the same arguments as in [Berestycki and Hamel 2002] to derive the Lipschitz continuity of the function $\lambda \mapsto \lambda\lfloor c_\lambda \rfloor$. If the functions c_λ admit a uniform mean then these arguments actually imply that $\lambda \mapsto \lambda\lfloor c_\lambda \rfloor$ is convex, but we do not know if this is true in general.

Lemma 3.3. *The functions $\lambda \mapsto \lfloor c_\lambda \rfloor$ and $\lambda \mapsto \lceil c_\lambda \rceil$ are locally uniformly Lipschitz continuous on* $(0, +\infty)$.

Proof. Fix $\Lambda > 0$ and $-\Lambda \le \lambda_0 \le \Lambda$. Let λ_1 be such that $|\lambda_1 - \lambda_0| = 2\Lambda$. For $j = 0, 1$, the function $v_j(x,t) := e^{-\lambda_j x\cdot e}\eta_{\lambda_j}(x,t)$ satisfies (17). Hence, setting $v_j = e^{w_j}$, we find that

$$\partial_t w_j - \mathrm{Tr}(AD^2 w_j) + q\cdot Dw_j = \mu + \mathrm{Tr}(ADw_j \otimes Dw_j), \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}.$$

For $\tau \in (0,1)$, the function $w := (1-\tau)w_0 + \tau w_1$ satisfies, for $x \in \mathbb{R}^N$, $t \in \mathbb{R}$,

$$\begin{aligned}\partial_t w - \mathrm{Tr}(AD^2 w) + q\cdot Dw &= \mu + \mathrm{Tr}\big(A((1-\tau)Dw_0 \otimes Dw_0 + \tau Dw_1 \otimes Dw_1)\big)\\ &\ge \mu + \mathrm{Tr}(ADw \otimes Dw).\end{aligned}$$

As a consequence, e^w is a supersolution of (17) and then, since

$$e^{w(x,t)} = e^{-((1-\tau)\lambda_0+\tau\lambda_1)x\cdot e}\eta_{\lambda_0}^{1-\tau}(x,t)\eta_{\lambda_1}^{\tau}(x,t),$$

the function $\eta_{\lambda_0}^{1-\tau}\eta_{\lambda_1}^{\tau}$ is a supersolution of (25) with $\lambda = \lambda_\tau := (1-\tau)\lambda_0 + \tau\lambda_1$. We can therefore apply the comparison principle between this function and η_{λ_τ} and derive, for $t \in \mathbb{R}$, $T > 0$,

$$\begin{aligned}\frac{\|\eta_{\lambda_\tau}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_\tau}(\cdot, t)\|_{L^\infty(\mathbb{R}^N)}} &\le \frac{\|\eta_{\lambda_0}^{1-\tau}\eta_{\lambda_1}^{\tau}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\min_{x\in\mathbb{R}^N}\eta_{\lambda_0}^{1-\tau}\eta_{\lambda_1}^{\tau}(x,t)}\\ &\le \bigg(\frac{\|\eta_{\lambda_0}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\min_{x\in\mathbb{R}^N}\eta_{\lambda_0}(x,t)}\bigg)^{1-\tau}\bigg(\frac{\|\eta_{\lambda_1}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\min_{x\in\mathbb{R}^N}\eta_{\lambda_1}(x,t)}\bigg)^{\tau}.\end{aligned}$$

Hence, using the inequality (28) for η_{λ_0} and η_{λ_1} (with the same C depending on Λ), we obtain

$$\frac{\|\eta_{\lambda_\tau}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_\tau}(\cdot, t)\|_{L^\infty(\mathbb{R}^N)}} \le C^{-1}\bigg(\frac{\|\eta_{\lambda_0}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_0}(\cdot, t)\|_{L^\infty(\mathbb{R}^N)}}\bigg)^{1-\tau}\bigg(\frac{\|\eta_{\lambda_1}(\cdot, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_1}(\cdot, t)\|_{L^\infty(\mathbb{R}^N)}}\bigg)^{\tau}. \tag{33}$$

Consider the function Γ defined by $\Gamma(\lambda) := \lambda\lfloor c_\lambda \rfloor$. It follows from (30) and (33) that

$$\Gamma(\lambda_\tau) \leq \lim_{T\to+\infty} \inf_{t\in\mathbb{R}} \frac{1}{T}\left((1-\tau)\ln \frac{\|\eta_{\lambda_0}(\cdot\,, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_0}(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}} + \tau \ln \frac{\|\eta_{\lambda_1}(\cdot\,, t+T)\|_{L^\infty(\mathbb{R}^N)}}{\|\eta_{\lambda_1}(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)}} \right).$$

If the $(c_\lambda)_{\lambda>0}$ admit a uniform mean, the above inequality and (30) imply that Γ is convex. Otherwise, we can only infer that

$$\Gamma(\lambda_\tau) \leq (1-\tau)\Gamma(\lambda_0) + \tau\lambda_1\lceil c_{\lambda_1}\rceil.$$

We have therefore shown that

$$\forall \tau \in (0,1), \quad \Gamma(\lambda_\tau) - \Gamma(\lambda_0) \leq \tau(\lambda_1\lceil c_{\lambda_1}\rceil - \lambda_0\lfloor c_{\lambda_0}\rfloor).$$

Thus, by (31) and (32) there exists a constant $K > 0$, depending on A, q, μ, such that

$$\forall \tau \in (0,1), \quad \Gamma(\lambda_\tau) - \Gamma(\lambda_0) \leq K(\Lambda^2+1)\tau.$$

This proves the Lipschitz continuity of Γ on $[-\Lambda, \Lambda]$, because $|\lambda_\tau - \lambda_0| = 2\Lambda\tau$, concluding the proof of the lemma.

The same arguments lead to the local Lipschitz continuity of $\lambda \mapsto \lceil c_\lambda \rceil$. □

3C. *Definition of the critical speed.* In order to define the critical speed c_*, we introduce the set

$$\Lambda := \{\lambda > 0 : \exists \bar{k} > 0,\ \forall 0 < k < \bar{k},\ \lfloor c_\lambda - c_{\lambda+k} \rfloor > 0\}. \tag{34}$$

Lemma 3.4. *There exists $\lambda_* > 0$ such that $\Lambda = (0, \lambda_*)$. Moreover, the function $\lambda \mapsto \lfloor c_\lambda \rfloor$ is decreasing on Λ.*

Proof. Fix $\lambda_0, \lambda_1 > 0$. For $\tau \in (0,1)$, we set $\lambda_\tau := (1-\tau)\lambda_0 + \tau\lambda_1$. Taking the natural log of (33) and recalling that $c_\lambda = S'_\lambda$ with S_λ satisfying (19) yields

$$\int_t^{t+T} [(1-\tau)\lambda_0 c_{\lambda_0} + \tau\lambda_1 c_{\lambda_1} - \lambda_\tau c_{\lambda_\tau}]\,ds \geq \ln C - 4\lambda_\tau\beta.$$

Hence,

$$\lambda_\tau \int_t^{t+T} (c_{\lambda_0} - c_{\lambda_\tau})\,ds \geq \tau\lambda_1 \int_t^{t+T} (c_{\lambda_0} - c_{\lambda_1})\,ds + \ln C - 4\lambda_\tau\beta.$$

Dividing both sides by T, taking the infimum over $t \in \mathbb{R}$ and then taking the limit as $T \to +\infty$, we derive

$$\forall \tau \in (0,1), \quad \lfloor c_{\lambda_0} - c_{\lambda_\tau} \rfloor \geq \tau \frac{\lambda_1}{\lambda_\tau} \lfloor c_{\lambda_0} - c_{\lambda_1} \rfloor. \tag{35}$$

If instead we divide by $-T$, we get

$$\forall \tau \in (0,1), \quad \lfloor c_{\lambda_\tau} - c_{\lambda_0} \rfloor \leq \tau \frac{\lambda_1}{\lambda_\tau} \lfloor c_{\lambda_1} - c_{\lambda_0} \rfloor. \tag{36}$$

Analogous estimates hold of course for the upper mean. The characterization of Λ follows from these inequalities, by suitable choices of λ_0, λ_1 and τ. We prove it in four steps.

Step 1: $\Lambda \neq \emptyset$. The first inequality in (31), together with (12), yields

$$\lim_{\lambda \to 0^+} \lfloor c_\lambda - c_1 \rfloor \geq \lim_{\lambda \to 0^+} \lfloor c_\lambda \rfloor - \lceil c_1 \rceil = +\infty.$$

Then there exists $0 < \lambda < 1$ such that $\lfloor c_\lambda - c_1 \rfloor > 0$. Applying (35) with $\lambda_0 = \lambda,\ \lambda_1 = 1$, we eventually infer that $\lfloor c_\lambda - c_{\lambda+k} \rfloor > 0$, for all $0 < k < 1 - \lambda$; that is, $\lambda \in \Lambda$.

Step 2: Λ *is bounded from above.* By (31) we obtain

$$\lim_{\lambda \to +\infty} \lfloor c_1 - c_\lambda \rfloor \leq \lfloor c_1 \rfloor - \lim_{\lambda \to +\infty} \lfloor c_\lambda \rfloor = -\infty.$$

Then there exists $\lambda' > 1$ such that, for $\lambda > \lambda'$, we have $\lfloor c_1 - c_\lambda \rfloor < 0$. Hence, for $k > 0$, applying (36) with $\lambda_0 = \lambda + k,\ \lambda_1 = 1$ and $\tau = k/(k + \lambda - 1)$, we derive

$$\lfloor c_\lambda - c_{\lambda+k} \rfloor \leq \frac{k}{(k+\lambda-1)\lambda} \lfloor c_1 - c_{\lambda+k} \rfloor < 0.$$

Namely, $\lambda \notin \Lambda$ and thus Λ is bounded from above by λ'.

Step 3: If $\lambda \in \Lambda$ *then* $(0, \lambda] \subset \Lambda$. Let $0 < \lambda' < \lambda$ and $k > 0$. Using first (35) and then (36) we get

$$\lfloor c_{\lambda'} - c_{\lambda'+k} \rfloor \geq \left(\frac{k}{k + \lambda - \lambda'} \right) \left(\frac{\lambda + k}{\lambda' + k} \right) \lfloor c_{\lambda'} - c_{\lambda+k} \rfloor \geq \left(\frac{\lambda + k}{\lambda' + k} \right) \frac{\lambda}{\lambda'} \lfloor c_\lambda - c_{\lambda+k} \rfloor.$$

Thus, $\lambda \in \Lambda$ implies $\lambda' \in \Lambda$.

Step 4: $\sup \Lambda \notin \Lambda$. Let $\lambda^* := \sup \Lambda$ and $k > 0$. For all $n \in \mathbb{N}$, there exists $0 < k_n < 1/n$ such that $\lfloor c_{\lambda^*+1/n} - c_{\lambda^*+1/n+k_n} \rfloor \leq 0$. For n large enough, we have that $1/n + k_n < k$ and then, by (35),

$$0 \geq \lfloor c_{\lambda^*+1/n} - c_{\lambda^*+1/n+k_n} \rfloor \geq \left(\frac{k_n}{k - 1/n} \right) \left(\frac{\lambda^* + k}{\lambda^* + 1/n + k_n} \right) \lfloor c_{\lambda^*+1/n} - c_{\lambda^*+k} \rfloor.$$

Hence,

$$\lfloor c_{\lambda^*} - c_{\lambda^*+k} \rfloor \leq \lfloor c_{\lambda^*+1/n} - c_{\lambda^*+k} \rfloor + \lceil c_{\lambda^*} - c_{\lambda^*+1/n} \rceil \leq \lceil c_{\lambda^*} - c_{\lambda^*+1/n} \rceil.$$

Using the analogue of (36) for the upper mean, we can control the latter term as follows:

$$\lceil c_{\lambda^*} - c_{\lambda^*+1/n} \rceil \leq \frac{1/n}{\lambda_* + 2/n} \lceil c_{\lambda^*/2} - c_{\lambda^*+1/n} \rceil \leq \frac{1/n}{\lambda_* + 2/n} (\lceil c_{\lambda^*/2} \rceil - \lfloor c_{\lambda^*+1/n} \rfloor),$$

which goes to 0 as $n \to \infty$ (recall that $\lambda \mapsto \lfloor c_\lambda \rfloor$ is continuous by Lemma 3.3). We eventually infer that $\lfloor c_{\lambda^*} - c_{\lambda^*+k} \rfloor \leq 0$; that is, $\lambda^* \notin \Lambda$.

It remains to show that $\lambda \mapsto \lfloor c_\lambda \rfloor$ is decreasing on Λ. Assume by way of contradiction that there are $0 < \lambda_1 < \lambda_2 < \lambda^*$ such that $\lfloor c_{\lambda_1} \rfloor \leq \lfloor c_{\lambda_2} \rfloor$. The function $\lambda \mapsto \lfloor c_\lambda \rfloor$, being continuous, attains its minimum on $[\lambda_1, \lambda_2]$ at some λ. Since $\lfloor c_{\lambda_1} \rfloor \leq \lfloor c_{\lambda_2} \rfloor$, we can assume that $\lambda \in [\lambda_1, \lambda_2)$. The definition of Λ implies that there exists $\lambda' \in (\lambda, \lambda_2)$ such that $\lfloor c_\lambda - c_{\lambda'} \rfloor > 0$. Hence, we obtain the following contradiction:

$$\lfloor c_{\lambda'} \rfloor \leq \lfloor c_\lambda \rfloor + \lceil c_{\lambda'} - c_\lambda \rceil = \lfloor c_\lambda \rfloor - \lfloor c_\lambda - c_{\lambda'} \rfloor < \lfloor c_\lambda \rfloor. \qquad \square$$

We are now in position to define the critical speed

$$c_* := \lfloor c_{\lambda_*} \rfloor, \tag{37}$$

where λ_* is given in Lemma 3.4.

Remark 1. When the terms in (1) are periodic in time, resuming from Section 2C, we know that the speeds $(c_\lambda)_{\lambda>0}$ are constant and satisfy $c_\lambda \equiv k(\lambda)/\lambda$, where $k(\lambda)$ is the principal eigenvalue of problem (20). Hence,

$$\lfloor c_\lambda - c_{\lambda+\kappa} \rfloor = \frac{k(\lambda)}{\lambda} - \frac{k(\lambda+\kappa)}{\lambda+\kappa}.$$

As $\lambda \mapsto k(\lambda)$ is strictly convex (see [Nadin 2009]) and, by (31),

$$\lim_{\lambda\to+\infty} \frac{k(\lambda)}{\lambda} = +\infty, \qquad \lim_{\lambda\to 0^+} k(\lambda) = \lim_{\lambda\to 0^+} \lambda c_\lambda \geq \Big\lfloor \min_{x\in\mathbb{R}^N} \mu(x,\cdot) \Big\rfloor > 0,$$

straightforward convexity arguments yield that λ_* is the unique minimizer of $\lambda \mapsto k(\lambda)/\lambda$. Therefore, $c_* = \min_{\lambda>0} k(\lambda)/\lambda$, which is known to be the minimal speed for pulsating traveling fronts (see [Nadin 2009]).

3D. ***Construction of a subsolution and conclusion of the proof.*** In order to prove Theorem 2.1, we introduce a family of functions $(\varphi_\lambda)_{\lambda>0}$ which play the role of the spatial-periodic principal eigenfunctions in the time-independent case. For $\lambda > 0$, let η_λ be the function given by Lemma 3.1, normalized by $\|\eta_\lambda(\cdot, 0)\|_{L^\infty(\mathbb{R}^N)} = 1$. We define

$$\varphi_\lambda(x,t) := e^{-\lambda S_\lambda(t)} \eta_\lambda(x,t).$$

By (19) and (28), there exist two positive constants C_λ, β such that

$$\forall x \in \mathbb{R}^N,\ t\in\mathbb{R}, \qquad C_\lambda \leq \varphi_\lambda(x,t) \leq e^{\lambda\beta}. \tag{38}$$

We will make use of the following key property of the least mean, provided by Lemma 3.2 of [Nadin and Rossi 2012]:

$$\forall g \in L^\infty(\mathbb{R}), \qquad \lfloor g \rfloor = \sup_{\sigma\in W^{1,\infty}(\mathbb{R})} \inf_{t\in\mathbb{R}} (\sigma' + g)(t). \tag{39}$$

Proof of Theorem 2.1. Fix $\gamma > c_*$. Since the function $\lambda \mapsto \lfloor c_\lambda \rfloor$ is continuous by Lemma 3.3 and tends to $+\infty$ as $\lambda \to 0^+$ by (31), and $\Lambda = (0, \lambda_*)$ by Lemma 3.4, there exists $\lambda \in \Lambda$ such that $\lfloor c_\lambda \rfloor = \gamma$. The function w defined by

$$w(x,t) := \min(e^{-\lambda x\cdot e} \eta_\lambda(x,t), 1)$$

is a generalized supersolution of (1).

In order to construct a subsolution, consider the constant ν in (10). By the definition of Λ, there exists $\lambda < \lambda' < (1+\nu)\lambda$ such that $\lfloor c_\lambda - c_{\lambda'} \rfloor > 0$. We then set $\psi(x,t) := e^{\sigma(t) - \lambda'(x\cdot e - S_\lambda(t) + S_{\lambda'}(t))} \eta_{\lambda'}(x,t)$, where $\sigma \in W^{1,\infty}(\mathbb{R})$ will be chosen later. We have that

$$\partial_t \psi - \operatorname{Tr}(A(x,t) D^2 \psi) + q(x,t)\cdot D\psi - \mu(x,t)\psi = \big(\sigma'(t) + \lambda'(c_\lambda(t) - c_{\lambda'}(t))\big)\psi.$$

Since $\lfloor\lambda'(c_\lambda - c_{\lambda'})\rfloor = \lambda'\lfloor c_\lambda - c_{\lambda'}\rfloor > 0$, by (39) we can choose $\sigma \in W^{1,\infty}(\mathbb{R})$ in such a way that $K := \inf_{\mathbb{R}}(\sigma' + \lambda'(c_\lambda - c_{\lambda'})) > 0$. Hence,

$$\partial_t\psi - \operatorname{Tr}(A(x,t)D^2\psi) + q(x,t)\cdot D\psi \geq (\mu(x,t)+K)\psi, \quad x\in\mathbb{R}^N,\ t\in\mathbb{R}.$$

We define

$$v(x,t) := e^{-\lambda x\cdot e}\eta_\lambda(x,t) - m\psi(x,t),$$

where m is a positive constant to be chosen. By computation,

$$e^{-\lambda x\cdot e}\eta_\lambda(x,t) - m\psi(x,t) = e^{-\lambda(x\cdot e - S_\lambda(t))}\big(\varphi_\lambda(x,t) - m\varphi_{\lambda'}(x,t)e^{\sigma(t)-(\lambda'-\lambda)(x\cdot e - S_\lambda(t))}\big).$$

Since φ_λ, $\varphi_{\lambda'}$ satisfy (38) and $\sigma\in L^\infty(\mathbb{R})$, it follows that, choosing m large enough, we have $v(x,t)\leq 0$ if $x\cdot e - S_\lambda(t)\leq 0$, and that v is less than $\delta\in(0,1]$, from (10), everywhere. If $v(x,t)>0$, and therefore $x\cdot e - S_\lambda(t) > 0$, we see that

$$\begin{aligned}
\partial_t v - \operatorname{Tr}(A(x,t)D^2 v) + q(x,t)\cdot Dv - \mu(x,t)v &\leq -mK\psi \\
&\leq -mK\psi\frac{v^{1+\nu}}{e^{-(1+\nu)\lambda x\cdot e}\eta_\lambda^{1+\nu}} \\
&= -mKv^{1+\nu}\frac{\varphi_{\lambda'}}{\varphi_\lambda^{1+\nu}}e^{\sigma(t)-(\lambda'-(1+\nu)\lambda)(x\cdot e - S_\lambda(t))} \\
&\leq -mKv^{1+\nu}C_{\lambda'}e^{-(1+\nu)\lambda\beta}\inf_{s\in\mathbb{R}} e^{\sigma(s)},
\end{aligned}$$

where, for the last inequality, we have used (38) and the fact that $\lambda' < (1+\nu)\lambda$. As a consequence, by hypothesis (10), for m sufficiently large, v is a subsolution of (1) in the set where it is positive.

Using again (38), one computes

$$\begin{aligned}
v(x+S_\lambda(t)e,t) &= e^{-\lambda x\cdot e}\big(\varphi_\lambda(x+S_\lambda(t)e,t) - m\varphi_{\lambda'}(x+S_\lambda(t)e,t)e^{\sigma(t)-(\lambda'-\lambda)x\cdot e}\big) \\
&\geq e^{-\lambda x\cdot e}(C_\lambda - mC_{\lambda'}e^{\lambda'\beta+\|\sigma\|_\infty-(\lambda'-\lambda)x\cdot e}).
\end{aligned}$$

Hence, taking R large enough, one has

$$\inf_{\substack{x\cdot e = R\\ t\in\mathbb{R}}} v(x+S_\lambda(t)e,t) \geq e^{-\lambda R}(C_\lambda - mC_{\lambda'}e^{\lambda'\beta+\|\sigma\|_\infty-(\lambda'-\lambda)R}) =: \omega\in(0,1).$$

Consequently, the function $\underline{v}$ defined by

$$\underline{v}(x,t) := \begin{cases} v(x,t) & \text{if } x\cdot e\geq S_\lambda(t)+R,\\ \max(\omega, v(x,t)) & \text{if } x\cdot e < S_\lambda(t)+R,\end{cases}$$

is continuous and, because of (8), it is a generalized subsolution of (1). Moreover, since $v\leq w$ and $w(x+S_\lambda(t)e,t)\geq e^{-\lambda R}C_\lambda > \omega$ if $x\cdot e < R$, one sees that $\underline{v}\leq w$. A solution $\underline{v}\leq u\leq w$ can therefore be obtained as the limit of (a subsequence of) the solutions $(u_n)_{n\in\mathbb{N}}$ of the problems

$$\begin{cases}\partial_t u_n - \operatorname{Tr}(A(x,t)D^2u_n) + q(x,t)\cdot Du_n = f(x,t,u_n), & x\in\mathbb{R}^N,\ t>-n\\ u_n(x,-n) = w(x,-n), & x\in\mathbb{R}^N.\end{cases}$$

The strong maximum principle yields $u > 0$. One further sees that

$$\lim_{x\cdot e\to+\infty} u\left(x+e\int_0^t c_\lambda(s)\,ds,\,t\right) \le \lim_{x\cdot e\to+\infty} w\left(x+e\int_0^t c_\lambda(s)\,ds,\,t\right) \le \lim_{x\cdot e\to+\infty} e^{-\lambda x\cdot e}\varphi_\lambda(x,t) = 0,$$

uniformly with respect to $t \in \mathbb{R}$. It remains to prove that

$$\lim_{x\cdot e\to-\infty} u\left(x+e\int_0^t c_\lambda(s)\,ds,\,t\right) = 1$$

holds uniformly with respect to $t \in \mathbb{R}$. Set

$$\vartheta := \lim_{r\to-\infty} \inf_{\substack{x\cdot e\le r\\ t\in\mathbb{R}}} u\left(x+e\int_0^t c_\lambda(s)\,ds,\,t\right).$$

Our aim is to show that $\vartheta = 1$. We know that $\vartheta \ge \omega > 0$, because $u(x,t) \ge \underline{v}(x,t) \ge \omega$ if $x\cdot e < S_\lambda(t)+R$. Let $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^N$ and $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ be such that

$$\lim_{n\to\infty} x_n\cdot e = -\infty, \qquad \lim_{n\to\infty} u\left(x_n+e\int_0^{t_n} c_\lambda(s)\,ds,\,t_n\right) = \vartheta.$$

For $n \in \mathbb{N}$, let $k_n \in \prod_{j=1}^N l_j\mathbb{Z}$ be such that $y_n := x_n + e\int_0^{t_n} c_\lambda(s)\,ds - k_n \in \prod_{j=1}^N [0, l_j)$ and define $v_n(x,t) := u(x+k_n, t+t_n)$. The functions $(v_n)_{n\in\mathbb{N}}$ are solutions of

$$\partial_t v_n - \mathrm{Tr}(A(x,t+t_n)D^2 v_n) + q(x,t+t_n)\cdot Dv_n = f(x,t+t_n,v_n), \quad x\in\mathbb{R}^N,\ t\in\mathbb{R}.$$

By parabolic estimates, one can show using the same types of arguments as in the proof of Lemma 3.1 that $(v_n)_{n\in\mathbb{N}}$ converges (up to subsequences) locally uniformly to some function v satisfying

$$\partial_t v - \mathrm{Tr}(\tilde{A}(x,t)D^2 v) + \tilde{q}(x,t)\cdot Dv = g(x,t) \ge 0, \quad x\in\mathbb{R}^N,\ t\in\mathbb{R},$$

where $\tilde{A}$ and $\tilde{q}$ are the strong limits in $L^\infty_{\mathrm{loc}}(\mathbb{R}^{N+1})$ and g is the weak limit in $L^p_{\mathrm{loc}}(\mathbb{R}^N\times\mathbb{R})$ of (a subsequence of) $A(x,t+t_n)$, $q(x,t+t_n)$ and $f(x,t+t_n,v_n(x,t))$ respectively, the inequality $g \ge 0$ coming from hypothesis (8). Furthermore, letting y be the limit of (a converging subsequence of) $(y_n)_{n\in\mathbb{N}}$, we find that $v(y,0) = \vartheta$ and

$$\forall x\in\mathbb{R}^N,\ t\in\mathbb{R}, \quad v(x,t) = \lim_{n\to\infty} u\left(x+x_n+e\int_0^{t_n} c_\lambda(s)\,ds - y_n,\, t+t_n\right) \ge \vartheta.$$

As a consequence, the strong maximum principle yields $v = \vartheta$ in $\mathbb{R}^N\times(-\infty,0]$. In particular, $g = 0$ in $\mathbb{R}^N\times(-\infty,0)$. Using the Lipschitz continuity of $f(x,t,\cdot)$, we then derive for all $(x,t)\in\mathbb{R}^N\times(-\infty,0)$,

$$\forall T>0, \quad 0 = \lim_{n\to+\infty} f(x,t+t_n,v_n(x,t)) = \lim_{n\to+\infty} f(x,t+t_n,\vartheta) \ge \inf_{(x,t)\in\mathbb{R}^{N+1}} f(x,t,\vartheta).$$

This, by (8), implies that either $\vartheta = 0$ or $\vartheta = 1$, whence $\vartheta = 1$ because $\vartheta \ge \omega > 0$. □

3E. ***A criterion for the existence of generalized transition waves in space-time general heterogeneous media.*** As already emphasized above, our proof holds in more general media, without assuming that the coefficients satisfy (11), that is, without the space periodicity assumption. We then need to assume that the linearized equation admits a family of solutions satisfying some global Harnack inequality. We conclude the existence part of the paper by stating such a result. We omit its proof since one only needs to check that the previous arguments still work.

Theorem 3.5. *In addition to (3)–(10), assume that there exists $\bar{\lambda} > 0$ such that, for all $\lambda \in (0, \bar{\lambda})$, there exists a Lipschitz-continuous time-global solution η_λ of*

$$\partial_t \eta_\lambda = \mathrm{Tr}(AD^2\eta_\lambda) - (q + 2\lambda Ae)D\eta_\lambda + (\mu + \lambda^2 eAe + \lambda q \cdot e)\eta_\lambda, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}$$

satisfying

$$\frac{1}{C}\|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)} e^{-CT} \le \eta_\lambda(x, t+T) \le C\|\eta_\lambda(\cdot\,, t)\|_{L^\infty(\mathbb{R}^N)} e^{CT},$$

for some $C = C(\lambda) > 0$ and for all $T > 0$, $(x, t) \in \mathbb{R}^{N+1}$.

Then there exists $\lambda_* \in (0, \bar{\lambda})$ such that, for all $\gamma > c_* := \lfloor S'_{\lambda_*} \rfloor$, where S_λ is a Lipschitz continuous function satisfying (19), there exists a generalized transition wave with speed $c_\lambda = S'_\lambda$ such that $\lfloor c_\lambda \rfloor = \gamma$.

4. Nonexistence result

Our aim is to find bounded subsolutions to the linearized problem

$$\partial_t u - \mathrm{Tr}(A(x,t)D^2u) + q(x,t)\cdot Du = \mu(x,t)u, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}, \tag{40}$$

in order to get a lower bound for the speed of traveling wave solutions. We recall that no spatial-periodic condition is now assumed. Looking for solutions of (40) in the form $u(x,t) = e^{-\lambda(x\cdot e - ct)}\phi(x,t)$, with λ and c constant, leads to the equation

$$(\mathcal{P}_\lambda + c\lambda)\phi = 0, \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}, \tag{41}$$

where $\mathcal{P}_\lambda$ is the linear parabolic operator defined by

$$\mathcal{P}_\lambda w := \partial_t w - \mathrm{Tr}(A(x,t)D^2w) + (q(x,t) - 2\lambda A(x,t)e)\cdot Dw - (\lambda^2 eA(x,t)e + \lambda q(x,t)\cdot e + \mu(x,t))w.$$

We consider the generalized principal eigenvalue introduced in [Berestycki and Nadin 2015]:

$$\kappa(\lambda) := \inf\bigl\{k \in \mathbb{R} : \exists\varphi,\ \inf\varphi > 0,\ \sup\varphi < \infty,\ \sup|D\varphi| < \infty,\ \mathcal{P}_\lambda\varphi \le k\varphi,\ \text{in } \mathbb{R}^N \times \mathbb{R}\bigr\}, \tag{42}$$

where the functions φ belong to $L^{N+1}_{\mathrm{loc}}(\mathbb{R}^{N+1})$, together with their derivatives ∂_t, D, D^2 (and therefore the differential inequalities are understood to hold almost everywhere). This is the minimal regularity required for the maximum principle to apply. See, e.g., [Lieberman 1996].

Taking $\varphi \equiv 1$ in the above definition we get, for $\lambda \in \mathbb{R}$,

$$\kappa(\lambda) \le -\underline{\alpha}\lambda^2 + \sup_{\mathbb{R}^{N+1}} |q||\lambda| - \inf_{\mathbb{R}^{N+1}} \mu. \tag{43}$$

We now derive a lower bound for $\kappa(\lambda)$. Assume by way of contradiction that there exists a function φ as in the definition of $\kappa(\lambda)$, associated with some k satisfying

$$k < -\bar{\alpha}\lambda^2 - \sup_{\mathbb{R}^{N+1}} |q||\lambda| - \sup_{\mathbb{R}^{N+1}} \mu.$$

For $\beta > 0$, the function $\psi(x,t) := e^{-\beta t}$ satisfies

$$\frac{\mathcal{P}_\lambda \psi}{\psi} \geq -\beta - \bar{\alpha}\lambda^2 - \sup_{\mathbb{R}^{N+1}} |q||\lambda| - \sup_{\mathbb{R}^{N+1}} \mu.$$

Hence, β can be chosen small enough in such a way that the latter term is larger than k; that is, $\mathcal{P}_\lambda \psi \geq k\psi$. The function ψ is larger than φ for t less than some t_0, whence $\psi \geq \varphi$ for all t by the comparison principle. It follows that $\varphi \to 0$ as $t \to +\infty$, which is impossible since φ is bounded from below away from 0. This shows that $\kappa(\lambda) > -\infty$.

We can now define c^* by setting

$$c^* := -\max_{\lambda > 0} \frac{\kappa(\lambda)}{\lambda}. \tag{44}$$

This definition is well posed if $\kappa(0) < 0$ because $\kappa(\lambda)/\lambda \to -\infty$ as $\lambda \to +\infty$ by (43), and we know from [Berestycki and Nadin 2015] that $\lambda \mapsto \kappa(\lambda)$ is Lipschitz-continuous.[3] Let us show that (12) implies that $\kappa(0) < 0$ and then that c^* is well defined and finite. Writing a positive function φ in the form $\varphi(t) := e^{-\sigma(t)}$, we see that

$$\mathcal{P}_0 \varphi = -(\sigma'(t) + \mu(x,t))\varphi \leq -\Big(\sigma'(t) + \inf_{x \in \mathbb{R}} \mu(x,t)\Big)\varphi.$$

Thus, (39) implies that, for given $\varepsilon > 0$, there exists $\sigma \in W^{1,\infty}(\mathbb{R})$ such that

$$\mathcal{P}_0 \varphi \leq -\bigg(\Big\lfloor \inf_{x \in \mathbb{R}^N} \mu(x,\cdot) \Big\rfloor - \varepsilon\bigg)\varphi.$$

Therefore, if (12) holds, taking $\varepsilon < \lfloor \min_{x \in \mathbb{R}^N} \mu(x,\cdot) \rfloor$ we derive $\kappa(0) < 0$.

The proof of Theorem 2.2 proceeds in two steps. In the following section we show that the average on $(0,+\infty)$ of the speed of a wave cannot be smaller than c^*. More precisely, we derive the following estimate.

Proposition 4.1. *Assume that (3)–(6) hold and that $\kappa(0) < 0$. Then, for any nonnegative supersolution u of (1) such that there is $c \in L^\infty(\mathbb{R})$ satisfying (2), it holds that*

$$\liminf_{t \to +\infty} \frac{1}{t} \int_0^t c(s)\,ds \geq c^* := -\max_{\lambda > 0} \frac{\kappa(\lambda)}{\lambda}.$$

In this statement, the notion of solution (including subsolution and supersolution) is understood as in the definition of $\kappa(\lambda)$: namely, u, $\partial_t u$, Du, $D^2 u \in L^{N+1}_{\mathrm{loc}}(\mathbb{R}^{N+1})$. Notice that the least mean of a function is in general smaller than the average on $(0,+\infty)$. In the last section, we establish a general property

[3] The coefficients are assumed to be Hölder continuous in [Berestycki and Nadin 2015], but one can check that it does not matter in the proof of continuity.

of the least mean that allows us to deduce Theorem 2.2 by applying Proposition 4.1 to suitable time translations of the original problem.

4A. *Lower bound on the mean speed for positive times.* We start by constructing subsolutions with a slightly varying exponential behavior as $x \cdot e \to \pm\infty$. These will then be used to build a generalized subsolution with an arbitrary modulation of the exponential behavior. The term "generalized subsolution" refers to a function that, in a neighborhood of each point, is obtained as the supremum of some family of subsolutions. Then, using the fact that the generalized subsolutions satisfy the maximum principle, we will be able to prove Proposition 4.1.

Lemma 4.2. *Let $c, \lambda \in \mathbb{R}$ be such that $\kappa(\lambda) + c\lambda < 0$. Then there exists $\varepsilon > 0$ and $M > 1$ such that, for any $z \in \mathbb{R}$, (40) admits a subsolution $\underline{v}$ satisfying*

$$\text{if } x \cdot e - ct \geq z, \quad \frac{1}{M} e^{-(\lambda+\varepsilon)(x\cdot e - ct)} \leq \underline{v}(x,t) \leq M e^{-(\lambda+\varepsilon)(x\cdot e-ct)}, \qquad \inf_{z-1<x\cdot e-ct<z} \underline{v}(x,t) > 0,$$

$$\text{if } x \cdot e - ct \leq z-1, \quad \frac{1}{M} e^{-(\lambda-\varepsilon)(x\cdot e - ct)} \leq \underline{v}(x,t) \leq M e^{-(\lambda-\varepsilon)(x\cdot e-ct)}.$$

Proof. By the definition of $\kappa(\lambda)$, there is a bounded function φ with positive infimum satisfying

$$\mathcal{P}_\lambda \varphi \leq k\varphi, \quad x \in \mathbb{R}^N,\ t > T,$$

for some $k < -c\lambda$. It follows that $v(x,t) := e^{-\lambda(x\cdot e-ct)}\varphi(x,t)$ is a subsolution of (40). Fix $z \in \mathbb{R}$ and consider a smooth function $\zeta : \mathbb{R} \to \mathbb{R}$ satisfying

$$\zeta = \lambda - \varepsilon \text{ in } (-\infty, z-1], \quad \zeta = \lambda + \varepsilon \text{ in } [z, +\infty), \quad 0 \leq \zeta' \leq 3\varepsilon, \quad |\zeta''| \leq h\varepsilon,$$

where $\varepsilon > 0$ has to be chosen and h is a universal constant. We define the function $\underline{v}$ by setting $\underline{v}(x,t) := e^{-(x\cdot e-ct)\zeta(x\cdot e-ct)}\varphi(x,t)$. Calling $\rho := x \cdot e - ct$, we find that

$$[\partial_t \underline{v} - a_{ij}(x,t)\partial_{ij}\underline{v} + q_i(x,t)\partial_i \underline{v} - \mu(x,t)\underline{v}]e^{\rho} \leq (\mathcal{P}_\zeta + c\zeta)\varphi + C[(1+\rho+\rho|\zeta| + \rho^2|\zeta'|)|\zeta'| + \rho|\zeta''|],$$

where ζ, ζ', ζ'' are evaluated at ρ, and C is a constant depending on N, c and the L^∞ norms of a_{ij}, q, μ, φ, $D\varphi$. The second term of the above right-hand side is bounded by $H(\varepsilon)$, for some continuous function H vanishing at 0. The first term satisfies

$$(\mathcal{P}_\zeta + c\zeta)\varphi \leq (\mathcal{P}_\lambda + c\lambda)\varphi + C((\zeta-\lambda) + |\zeta^2 - \lambda^2|) \leq (k + c\lambda)\varphi + C(\varepsilon + 2|\lambda|\varepsilon + \varepsilon^2).$$

We thus derive

$$\partial_t \underline{v} - a_{ij}(x,t)\partial_{ij}\underline{v} + q_i(x,t)\partial_i \underline{v} - \mu(x,t)\underline{v} \leq e^{\rho}[(k+c\lambda)\varphi + C\varepsilon(1 + 2|\lambda| + \varepsilon^2) + H(\varepsilon)].$$

Since $k < -c\lambda$ and $\inf \varphi > 0$, we can choose ε small enough in such a way that $\underline{v}$ is a subsolution of (40). □

Lemma 4.3. *Let $\underline{\lambda}, \bar{\lambda}, c \in \mathbb{R}$ satisfy $\underline{\lambda} < \bar{\lambda}$ and*

$$\max_{\lambda \in [\underline{\lambda}, \bar{\lambda}]} (\kappa(\lambda) + c\lambda) < 0.$$

Then there exists a generalized, bounded subsolution $\underline{v}$ *of* (40) *satisfying*

$$\lim_{r\to-\infty}\sup_{x\cdot e-ct<r}\underline{v}(x,t)e^{\underline{\lambda}(x\cdot e-ct)}=0,\qquad \lim_{r\to+\infty}\sup_{x\cdot e-ct>r}\underline{v}(x,t)e^{\bar{\lambda}(x\cdot e-ct)}=0,$$

and

$$\forall r_1<r_2,\qquad \inf_{r_1<x\cdot e-ct<r_2}\underline{v}(x,t)>0. \tag{45}$$

Proof. For $\lambda\in[\underline{\lambda},\bar{\lambda}]$, let ε_λ, M_λ be the constants given by Lemma 4.2 associated with c and λ. Call I_λ the interval $(\lambda-\varepsilon_\lambda,\lambda+\varepsilon_\lambda)$. The family $(I_\lambda)_{\underline{\lambda}\le\lambda\le\bar{\lambda}}$ is an open covering of $[\underline{\lambda},\bar{\lambda}]$. Let $(I_{\lambda_i})_{i=1,\dots,n}$ be a finite subcovering and set for short $\varepsilon_i:=\varepsilon_{\lambda_i}$, $M_i:=M_{\lambda_i}$. Up to rearranging the indices and extracting another subcovering if need be, we can assume that

$$\forall i=1,\dots,n-1,\qquad \lambda_{i+1}-\varepsilon_{i+1}<\lambda_i-\varepsilon_i<\lambda_{i+1}+\varepsilon_{i+1}<\lambda_i+\varepsilon_i.$$

Let v_1 be the subsolution of (40) given by Lemma 4.2 associated with $\lambda=\lambda_1$ and $z=0$. Set $z_1:=0$, $k_1:=1$ and

$$k_2:=\frac{e^{(\lambda_2+\varepsilon_2-(\lambda_1-\varepsilon_1))(z_1-1)}}{M_1M_2}.$$

Consider then the subsolution v_2 associated with $\lambda=\lambda_2$ and z equal to some value $z_2<z_1-1$ to be chosen. We have that

$$\text{if } x\cdot e-ct=z_1-1,\qquad \frac{v_1(x,t)}{v_2(x,t)}\ge\frac{k_2}{k_1}$$
$$\text{if } x\cdot e-ct=z_2,\qquad \frac{k_1v_1(x,t)}{k_2v_2(x,t)}\le(M_1M_2)^2e^{(\lambda_2+\varepsilon_2-(\lambda_1-\varepsilon_1))(z_2-z_1+1)}.$$

Since $\lambda_2+\varepsilon_2>\lambda_1-\varepsilon_1$, we can choose z_2 in such a way that the latter term is less than 1. By a recursive argument, we find some constants $(z_i)_{i=1,\dots,n}$ satisfying $z_n<z_{n-1}-1<\cdots<z_1-1=-1$, such that the family of subsolutions $(v_i)_{i=1,\dots,n}$ given by Lemma 4.2 associated with the $(\lambda_i)_{i=1,\dots,n}$ and $(z_i)_{i=1,\dots,n}$ satisfies, for some positive $(k_i)_{i=1,\dots,n}$,

$$\forall i=1,\dots,n-1,\quad k_{i+1}v_{i+1}\le k_iv_i\ \text{ if } x\cdot e-ct=z_i-1,\quad k_{i+1}v_{i+1}\ge k_iv_i\ \text{ if } x\cdot e-ct=z_{i+1}.$$

The function $\underline{v}$, defined by

$$\underline{v}(x,t):=\begin{cases}v_1(x,t) & \text{if } x\cdot e-ct\ge z_1,\\ \max(k_iv_i(x,t),k_{i+1}v_{i+1}(x,t)) & \text{if } z_{i+1}\le x\cdot e-ct<z_i,\\ k_nv_n(x,t) & \text{if } x\cdot e-ct<z_n,\end{cases}$$

is a generalized subsolution of (40) satisfying the desired properties. □

Proof of Proposition 4.1. Let u, c be as in the statement of the proposition, and define $\phi(x,t):=u(x+e\int_0^t c(s)\,ds,t)$. Since $\phi(x,t)\to1$ as $x\cdot e\to-\infty$, uniformly with respect to $t\in\mathbb{R}$, one can find $\rho\in\mathbb{R}$ such that

$$\inf_{\substack{x\cdot e<\rho\\ t\in\mathbb{R}}}\phi(x,t)>0.$$

We now make use of Lemma 3.1 in [Rossi and Ryzhik 2014], which, under the above condition, establishes a lower bound for the exponential decay of an entire supersolution ϕ of a linear parabolic equation (notice that the differential inequality for ϕ can be written in linear form with a bounded zero order term: $f(x,t,\phi) = [f(x,t,\phi)/\phi]\phi$). The result of [Rossi and Ryzhik 2014] implies the existence of a positive constant λ_0 such that

$$\inf_{\substack{x\cdot e>\rho-1\\ t\in\mathbb{R}}} \phi(x,t)e^{\lambda_0 x\cdot e} > 0.$$

By the definition of c^*, the hypotheses of Lemma 4.3 are fulfilled with $\underline{\lambda} = 0$, $\bar{\lambda} = \lambda_0$ and $c = c^* - \varepsilon$, for any given $\varepsilon > 0$. This is also true if one penalizes the nonlinear term $f(x,t,u)$ by subtracting δu, with δ small enough, since this just raises the principal eigenvalues $\kappa(\lambda)$ by δ. Therefore, Lemma 4.3 provides a function $\underline{v}$ such that, for $h > 0$ small enough, $h\underline{v}$ is a subsolution of (1). We choose h in such a way that, together with the above property, $h\underline{v}(x,0) < u(x,0)$. This can be done, due to the lower bounds of $u(x,0) = \phi(x,0)$, because $\underline{v}$ is bounded and decays faster than $e^{-\lambda_0 x\cdot e}$ as $x\cdot e \to +\infty$. Applying the parabolic comparison principle we eventually infer that $h\underline{v} < u$ for all $x \in \mathbb{R}^N$, $t \geq 0$. It follows that u satisfies (45) with $c = c^* - \varepsilon$ for $t > 0$. We derive, in particular,

$$0 < \inf_{t>0} u((c^*-\varepsilon)te, t) = \inf_{t>0} u\left(\left((c^*-\varepsilon)t - \int_0^t c(s)\,ds\right)e + e\int_0^t c(s)\,ds, t\right),$$

which, in virtue of the second condition in (2), implies that

$$\limsup_{t\to+\infty}\left((c^*-\varepsilon)t - \int_0^t c(s)\,ds\right) < +\infty.$$

This concludes the proof due to the arbitrariness of ε. □

4B. *Property of the least mean and proof of Theorem 2.2.* Roughly speaking, the least mean of a function is the infimum of its averages in sufficiently large intervals. We show that, in some sense, this infimum is achieved up to replacing the function with an element of its *ω-limit* set. The ω-limit (in the L^∞ weak-$*$ topology) of a function g, denoted by ω_g, is the set of functions obtained as L^∞ weak-$*$ limits of translations of g.

Proposition 4.4. *Let $g \in L^\infty(\mathbb{R})$ and let ω_g denote its ω-limit set (in the L^∞ weak-$*$ topology). Then*

$$\lfloor g\rfloor = \min_{\tilde{g}\in\omega_g}\left(\lim_{t\to+\infty}\frac{1}{t}\int_0^t \tilde{g}(s)\,ds\right).$$

Proof. We can assume without loss of generality that $\lfloor g\rfloor = 0$. Clearly, any $\tilde{g} \in \omega_g$ satisfies $\lfloor\tilde{g}\rfloor \geq \lfloor g\rfloor$, whence

$$\liminf_{t\to+\infty}\frac{1}{t}\int_0^t \tilde{g}(s)\,ds \geq \lfloor\tilde{g}\rfloor \geq \lfloor g\rfloor = 0.$$

Our aim is to find a function $\tilde{g} \in \omega_g$ satisfying

$$\limsup_{t\to+\infty}\frac{1}{t}\int_0^t \tilde{g}(s)\,ds \leq 0. \tag{46}$$

We claim that, for any $n \in \mathbb{N}$, there exists $t_n \in \mathbb{N}$ such that

$$\forall j = 1, \dots, n, \quad n \int_{t_n}^{t_n+j} g(s)\, ds \le j.$$

Assume by way of contradiction that this property fails for some $n \in \mathbb{N}$. By the definition of least mean, for $K \in \mathbb{N}$ large enough, there is $\tau \in \mathbb{R}$ such that

$$\frac{1}{K} \int_{\tau}^{\tau+Kn} g(s)\, ds < \frac{1}{2}.$$

On the other hand, there is $j \in \{1, \dots, n\}$ such that $n \int_\tau^{\tau+j} g(s)\, ds > j$. Then, there is $h \in \{1, \dots, n\}$ such that $n \int_{\tau+j}^{\tau+j+h} g(s)\, ds > h$, and hence $n \int_\tau^{\tau+j+h} g(s)\, ds > j + h$. We repeat this argument until we find $k \in \{1, \dots, n\}$ such that $n \int_\tau^{\tau+Kn+k} g(s)\, ds > Kn + k$. From this we deduce that

$$\int_\tau^{\tau+Kn} g(s)\, ds > K + \frac{k}{n} - \int_{\tau+Kn}^{\tau+Kn+k} g(s)\, ds > K - n\|g\|_{L^\infty(\mathbb{R})}.$$

A contradiction follows taking $K > 2n\|g\|_{L^\infty(\mathbb{R})}$, and the claim is proved. The L^∞ weak-$*$ limit $\tilde{g}$ as $n \to \infty$ of (a subsequence of) $g(\cdot + t_n)$ satisfies the desired property. Indeed,

$$\forall j \in \mathbb{N}, \quad \int_0^j \tilde{g}(s)\, ds = \lim_{n\to\infty} \int_{t_n}^{t_n+j} g(s)\, ds = 0,$$

from which (46) follows since $\tilde{g}$ is bounded. □

Proof of Theorem 2.2. Let u be a generalized transition wave with speed c. Proposition 4.4 yields that there exists $\tilde{c} \in \omega_c$ such that

$$\lfloor c \rfloor = \lim_{T\to+\infty} \frac{1}{T} \int_0^T \tilde{c}(s)\, ds. \tag{47}$$

The definition of ω_c gives a sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that $c(\cdot + t_n) \rightharpoonup \tilde{c}$ as $n \to +\infty$ for the L^∞ weak-$*$ topology. For $n \in \mathbb{N}$, consider the functions

$$A_n(x,t) := A\Big(x + e\int_0^{t_n} c(s)\, ds, t + t_n\Big), \quad q_n(x,t) := q\Big(x + e\int_0^{t_n} c(s)\, ds, t + t_n\Big),$$
$$\mu_n(x,t) := \mu\Big(x + e\int_0^{t_n} c(s)\, ds, t + t_n\Big), \quad u_n(x,t) := u\Big(x + e\int_0^{t_n} c(s)\, ds, t + t_n\Big).$$

For any $\varepsilon \in (0,1)$ there exists $m \in (0,1)$ such that

$$\forall (x,t) \in \mathbb{R}^{N+1},\ u \in [0,1], \quad f(x,t,u) \ge (\mu(x,t) - \varepsilon) u (m - u).$$

It follows that the u_n satisfy

$$\partial_t u_n - \mathrm{Tr}(A_n(x,t) D^2 u_n) + q_n(x,t) D u_n \ge (\mu_n(x,t) - \varepsilon) u_n (m - u_n), \quad x \in \mathbb{R}^N,\ t \in \mathbb{R}.$$

On the other hand, the L^p parabolic interior estimates ensure that the sequences $(\partial_t u_n)_{n\in\mathbb{N}}$, $(D u_n)_{n\in\mathbb{N}}$, $(D^2 u_n)_{n\in\mathbb{N}}$ are bounded in $L^p(Q)$ for all $p \in (1,\infty)$ and $Q \Subset \mathbb{R}^{N+1}$. Hence, by the embedding theorem, $(u_n)_{n\in\mathbb{N}}$ converges (up to subsequences) locally uniformly in $\mathbb{R}^{N+1}$ to some function $\tilde{u}$, and the derivatives ∂_t, D, D^2 of the $(u_n)_{n\in\mathbb{N}}$ weakly converge to those of $\tilde{u}$ in $L^p_{\mathrm{loc}}(\mathbb{R}^{N+1})$. Therefore, letting $\tilde{A}$, $\tilde{q}$ be the locally uniform limits of (subsequences of) $(A_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$ and $\tilde{\mu}$ be the L^∞ weak-$*$ limit of (a subsequence of) $(\mu_n)_{n\in\mathbb{N}}$, we infer that

$$\partial_t\tilde{u} - \mathrm{Tr}(\tilde{A}(x,t)D^2\tilde{u}) + \tilde{q}(x,t)D\tilde{u} \geq (\tilde{\mu}(x,t)-\varepsilon)\tilde{u}(m-\tilde{u}), \quad x\in\mathbb{R}^N,\ t\in\mathbb{R}.$$

Hence, $\tilde{u}$ is a supersolution of an equation of the type (1) whose terms satisfy (3)–(5) and (6). Moreover, it is easily derived from the definition of the speed c and the L^∞ weak-$*$ convergence to $\tilde{c}$, that $\tilde{u}$ satisfies (2) with c replaced by $\tilde{c}$, uniformly with respect to $t\in\mathbb{R}$. In order to apply Proposition 4.1 to the function $\tilde{u}$, we need to show that $\tilde{\kappa}(0)<0$, where $\lambda\mapsto\tilde{\kappa}(\lambda)$ is defined like $\lambda\mapsto\kappa(\lambda)$, but with $\tilde{A}$, $\tilde{q}$, $\tilde{\mu}-\varepsilon$ in place of A, q, μ respectively. Namely, the $\kappa(\lambda)$ are the principal eigenvalues in the sense of (42) for the operators $\tilde{\mathcal{P}}_\lambda$ defined as follows:

$$\tilde{\mathcal{P}}_\lambda w := \partial_t w - \mathrm{Tr}(\tilde{A}(x,t)D^2w) + (\tilde{q}(x,t)-2\lambda\tilde{A}(x,t)e)\cdot Dw - (\lambda^2 e\tilde{A}(x,t)e + \lambda\tilde{q}(x,t)\cdot e + \tilde{\mu}(x,t)-\varepsilon)w.$$

This will be achieved by showing that

$$\forall\lambda>0,\quad \tilde{\kappa}(\lambda)\leq\kappa(\lambda)+\varepsilon, \tag{48}$$

whence $\tilde{\kappa}(0)<0$ as soon as $\varepsilon<-\kappa(0)$ (recall that $\kappa(0)<0$ by (12)). Let us postpone for a moment the proof of (48). Applying Proposition 4.1 to $\tilde{u}$ yields

$$\liminf_{t\to+\infty}\frac{1}{t}\int_0^t \tilde{c}(s)\,ds \geq -\max_{\lambda>0}\frac{\tilde{\kappa}(\lambda)}{\lambda} = -\frac{\tilde{\kappa}(\hat{\lambda})}{\hat{\lambda}},$$

for some $\hat{\lambda}>0$. In virtue of (47) and (48), from this inequality we deduce

$$\lfloor c\rfloor \geq -\frac{\kappa(\hat{\lambda})+\varepsilon}{\hat{\lambda}},$$

from which $\lfloor c\rfloor\geq c^*$ follows by the arbitrariness of ε.

It remains to prove (48). Let $k>\kappa(\lambda)$. By definition (43) there exists φ such that $\inf\varphi>0$ and $\varphi, D\varphi\in L^\infty(\mathbb{R}^N\times\mathbb{R})$ and $\mathcal{P}_\lambda\varphi\leq k\varphi$ in $\mathbb{R}^N\times\mathbb{R}$. We would like to perform on φ the same limit of translations as done before to obtain $\tilde{u}$ from u. This would yield a function $\tilde{\varphi}$ satisfying $\tilde{\mathcal{P}}_\lambda\tilde{\varphi}\leq(k+\varepsilon)\tilde{\varphi}$. But this argument requires the L^p_{loc} estimates of the derivatives ∂_t, D, D^2 of the translated of φ, which are not available because φ is a subsolution and not a solution of an equation. However, it is possible to replace φ with a solution of a semilinear equation of the type $\mathcal{P}_\lambda w = g(w)$ in $\mathbb{R}^N\times\mathbb{R}$, with g smooth and such that $g(w)\leq(k+\varepsilon)w$, which satisfies the same properties as φ, as well as the desired additional regularity properties. This is done in the proof of Theorem A.1 of [Rossi and Ryzhik 2014], whose arguments can be exactly repeated here. We can therefore apply the translation argument that provides a function $\tilde{\varphi}$ such that $\tilde{\mathcal{P}}_\lambda\tilde{\varphi}\leq(k+\varepsilon)\tilde{\varphi}$. Moreover, $\inf\tilde{\varphi}>0$ and $\sup\tilde{\varphi}<\infty$. In order to be able to use $\tilde{\varphi}$ in the definition of $\tilde{\kappa}(\lambda)$ and derive $\tilde{\kappa}(\lambda)\leq k+\varepsilon$, we only need to have that $\sup|D\tilde{\varphi}|<\infty$. This property does not follow automatically from the L^p estimates and the embedding theorem as in the elliptic case treated in [Rossi and Ryzhik 2014]. This is the reason why we need the extra assumption (13) on A. Indeed, we use Theorem 1.4 of [Porretta and Priola 2013] with, using the same notations as in [Porretta and Priola 2013], F the nonlinear operator associated with equation $\mathcal{P}_\lambda w = g(w)$. Hypothesis 1.2 of [Porretta and Priola 2013] is satisfied since A satisfies (13), q is bounded and $f = f(x,t,u)$ is bounded with respect to $(x,t,u) \in \mathbb{R}^N \times \mathbb{R} \times [0,1]$, and Hypothesis 1.3 is satisfied with $\varphi(x,t) := e^{Mt}(1+|x|^2)$ and M large enough. Hence, we get a uniform L^∞ bound on Dw, where w is the solution of $\mathcal{P}_\lambda w = g(w)$. Using w instead of φ, we get that this bound is inherited by $\tilde{\varphi}$ and we therefore deduce $\tilde{\kappa}(\lambda) \le k+\varepsilon$. As $k > \kappa(\lambda)$ is arbitrary, we eventually get (48). □

References

[Alikakos et al. 1999] N. D. Alikakos, P. W. Bates, and X. Chen, "Periodic traveling waves and locating oscillating patterns in multidimensional domains", *Trans. Amer. Math. Soc.* **351**:7 (1999), 2777–2805. MR 99j:35101 Zbl 0929.35067

[Aronson and Weinberger 1978] D. G. Aronson and H. F. Weinberger, "Multidimensional nonlinear diffusion arising in population genetics", *Adv. in Math.* **30**:1 (1978), 33–76. MR 80a:35013 Zbl 0407.92014

[Berestycki and Hamel 2002] H. Berestycki and F. Hamel, "Front propagation in periodic excitable media", *Comm. Pure Appl. Math.* **55**:8 (2002), 949–1032. MR 2003d:35139 Zbl 1024.37054

[Berestycki and Hamel 2007] H. Berestycki and F. Hamel, "Generalized travelling waves for reaction-diffusion equations", pp. 101–123 in *Perspectives in nonlinear partial differential equations*, edited by H. Berestycki et al., Contemp. Math. **446**, Amer. Math. Soc., Providence, RI, 2007. MR 2009k:35135 Zbl 1200.35169

[Berestycki and Hamel 2012] H. Berestycki and F. Hamel, "Generalized transition waves and their properties", *Comm. Pure Appl. Math.* **65**:5 (2012), 592–648. MR 2898886 Zbl 1248.35039

[Berestycki and Nadin 2015] H. Berestycki and G. Nadin, "Asymptotic spreading for general heterogeneous Fisher-KPP type equations", preprint, 2015, available at https://edutice.archives-ouvertes.fr/hal-01171334/document.

[Berestycki and Nirenberg 1992] H. Berestycki and L. Nirenberg, "Travelling fronts in cylinders", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **9**:5 (1992), 497–572. MR 93k:35019 Zbl 0799.35073

[Berestycki et al. 2005] H. Berestycki, F. Hamel, and L. Roques, "Analysis of the periodically fragmented environment model, II: Biological invasions and pulsating travelling fronts", *J. Math. Pures Appl.* (9) **84**:8 (2005), 1101–1146. MR 2006d:35123 Zbl 1083.92036

[Berestycki et al. 2008] H. Berestycki, F. Hamel, and G. Nadin, "Asymptotic spreading in heterogeneous diffusive excitable media", *J. Funct. Anal.* **255**:9 (2008), 2146–2189. MR 2009k:35136 Zbl 1234.35033

[Fisher 1937] R. A. Fisher, "The wave of advance of advantageous genes", *Ann. Eugenics* **7**:4 (1937), 355–369. Zbl 63.1111.04

[Kolmogorov et al. 1937] A. Kolmogorov, I. Petrovsky, and N. Piskunov, "Etude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique", *Bulletin Université d'Etat à Moscou* (*Bjul. Moskowskogo Gos. Univ*) **1**:6 (1937), 1–26. Zbl 0018.32106

[Lieberman 1996] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific Publishing Co., River Edge, NJ, 1996. MR 98k:35003 Zbl 0884.35001

[Mellet et al. 2009] A. Mellet, J. Nolen, J.-M. Roquejoffre, and L. Ryzhik, "Stability of generalized transition fronts", *Comm. Partial Differential Equations* **34**:4-6 (2009), 521–552. MR 2010d:35185 Zbl 1173.35021

[Mellet et al. 2010] A. Mellet, J.-M. Roquejoffre, and Y. Sire, "Generalized fronts for one-dimensional reaction-diffusion equations", *Discrete Contin. Dyn. Syst.* **26**:1 (2010), 303–312. MR 2011b:35238 Zbl 1180.35294

[Nadin 2009] G. Nadin, "Traveling fronts in space-time periodic media", *J. Math. Pures Appl.* (9) **92**:3 (2009), 232–262. MR 2010i:35193 Zbl 1182.35074

[Nadin 2014] G. Nadin, "Critical travelling waves for general heterogeneous reaction-diffusion", preprint, 2014, available at https://www.ljll.math.upmc.fr/ nadin/criticalTWrevised.pdf.

[Nadin and Rossi 2012] G. Nadin and L. Rossi, "Propagation phenomena for time heterogeneous KPP reaction-diffusion equations", *J. Math. Pures Appl.* (9) **98**:6 (2012), 633–653. MR 2994696 Zbl 06115427

[Nolen and Ryzhik 2009] J. Nolen and L. Ryzhik, "Traveling waves in a one-dimensional heterogeneous medium", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:3 (2009), 1021–1047. MR 2010d:35201 Zbl 1178.35205

[Nolen et al. 2005] J. Nolen, M. Rudd, and J. Xin, "Existence of KPP fronts in spatially-temporally periodic advection and variational principle for propagation speeds", *Dyn. Partial Differ. Equ.* **2**:1 (2005), 1–24. MR 2006h:35129 Zbl 1179.35166

[Nolen et al. 2012] J. Nolen, J.-M. Roquejoffre, L. Ryzhik, and A. Zlatoš, "Existence and non-existence of Fisher-KPP transition fronts", *Arch. Ration. Mech. Anal.* **203**:1 (2012), 217–246. MR 2864411 Zbl 1267.35108

[Porretta and Priola 2013] A. Porretta and E. Priola, "Global Lipschitz regularizing effects for linear and nonlinear parabolic equations", *J. Math. Pures Appl.* (9) **100**:5 (2013), 633–686. MR 3115828 Zbl 06448859

[Rossi and Ryzhik 2014] L. Rossi and L. Ryzhik, "Transition waves for a class of space-time dependent monostable equations", *Commun. Math. Sci.* **12**:5 (2014), 879–900. MR 3187783 Zbl 1305.35092

[Shen 1999] W. Shen, "Travelling waves in time almost periodic structures governed by bistable nonlinearities, II: Existence", *J. Differential Equations* **159**:1 (1999), 55–101. MR 2001k:35158 Zbl 0939.35015

[Shen 2006] W. Shen, "Traveling waves in time dependent bistable equations", *Differential Integral Equations* **19**:3 (2006), 241–278. MR 2007a:35072 Zbl 1212.35220

[Shen 2011a] W. Shen, "Existence of generalized traveling waves in time recurrent and space periodic monostable equations", *J. Appl. Anal. Comput.* **1**:1 (2011), 69–93. MR 2881848 Zbl 1304.35351

[Shen 2011b] W. Shen, "Existence, uniqueness, and stability of generalized traveling waves in time dependent monostable equations", *J. Dynam. Differential Equations* **23**:1 (2011), 1–44. MR 2012d:35177 Zbl 1223.35103

[Xin 1992] J. X. Xin, "Existence of planar flame fronts in convective-diffusive periodic media", *Arch. Rational Mech. Anal.* **121**:3 (1992), 205–233. MR 93m:35110 Zbl 0764.76074

[Zlatoš 2013] A. Zlatoš, "Generalized traveling waves in disordered media: existence, uniqueness, and stability", *Arch. Ration. Mech. Anal.* **208**:2 (2013), 447–480. MR 3035984 Zbl 1270.35173

Received 24 Mar 2014. Revised 15 Mar 2015. Accepted 11 May 2015.

GRÉGOIRE NADIN: gregoire.nadin@upmc.fr
Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France

LUCA ROSSI: lucar@math.unipd.it
Dipartimento di Matematica, Università di Padova, via Trieste 63, I-35121 Padova, Italy

ANALYSIS AND PDE
Vol. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1379

CHARACTERISATION OF THE ENERGY OF GAUSSIAN BEAMS ON LORENTZIAN MANIFOLDS: WITH APPLICATIONS TO BLACK HOLE SPACETIMES

JAN SBIERSKI

It is known that, using the Gaussian beam approximation, one can show that there exist solutions of the wave equation on a general globally hyperbolic Lorentzian manifold whose energy is localised along a given null geodesic for a finite, but arbitrarily long, time. We show that the energy of such a localised solution is determined by the energy of the underlying null geodesic. This result opens the door to various applications of Gaussian beams on Lorentzian manifolds that *do not admit a globally timelike Killing vector field*. In particular, we show that trapping in the exterior of Kerr or at the horizon of an extremal Reissner–Nordström black hole necessarily leads to a "loss of derivative" in a local energy decay statement. We also demonstrate the obstruction formed by the red-shift effect at the event horizon of a Schwarzschild black hole to scattering constructions from the future (where the red-shift turns into a blue-shift): we construct solutions to the backwards problem whose energies grow exponentially for a finite, but arbitrarily long, time. Finally, we give a simple mathematical realisation of the heuristics for the blue-shift effect near the Cauchy horizon of subextremal and extremal black holes: we construct a sequence of solutions to the wave equation whose initial energies are uniformly bounded, whereas the energy near the Cauchy horizon goes to infinity.

MSC2010: primary 35R01; secondary 83C57.
Keywords: Gaussian beams, characterisation of energy, spacetime.

1. Introduction

Part I of this paper is concerned with the study of the temporal behaviour of Gaussian beams on *general* globally hyperbolic Lorentzian manifolds. Here, a Gaussian beam is a highly oscillatory wave packet of the form

$$\tilde{u}_\lambda = \frac{1}{\sqrt{E(\lambda, a, \phi)}} \cdot a \cdot e^{i\lambda\phi},$$

where $E(\lambda, a, \phi)$ is a renormalisation factor keeping the initial energy of $\tilde{u}_\lambda$ independent of $\lambda \in \mathbb{R}^+$, and the complex-valued functions a and ϕ are chosen in such a way that for $\lambda \gg 0$ the Gaussian beam $\tilde{u}_\lambda$ is an approximate solution to the wave equation on the underlying Lorentzian manifold (M, g). The failure of $\tilde{u}_\lambda$ being an actual solution to the wave equation

$$\Box_g u = 0 \tag{1.1}$$

is measured in terms of an energy norm — and this error can be made arbitrarily small up to a finite, but arbitrarily long, time, by choosing λ large enough. The construction of the functions a and ϕ allows for restricting the support of a to a small neighbourhood of a given null geodesic. Thus, one can infer from $\tilde{u}_\lambda$ being an approximate solution with respect to some energy norm that:[1]

> There exist *actual* solutions of the wave equation (1.1) whose "energy" is localised along a given null geodesic up to some finite, but arbitrarily long, time. (1.2)

This is, roughly, the state of the art knowledge of Gaussian beams (see, for instance, [Ralston 1982]).

The main new result of Part I is to provide a geometric characterisation of the temporal behaviour of the localised energy of a Gaussian beam. More precisely, given a timelike vector field N (with respect to which we measure the energy) and a Gaussian beam $\tilde{u}_\lambda$ supported in a small neighbourhood of an affinely parametrised null geodesic γ, we show in Theorem 4.1 that

$$\int_{\Sigma_\tau} J^N(\tilde{u}_\lambda) \cdot n_{\Sigma_\tau} \approx -g(N, \dot{\gamma})\big|_{\mathrm{Im}(\gamma)\cap\Sigma_\tau} \tag{1.3}$$

holds up to some finite time T. Here, we consider a foliation of the Lorentzian manifold (M, g) by spacelike slices Σ_τ, $J^N(\tilde{u}_\lambda)$ denotes the contraction of the stress–energy tensor[2] of $\tilde{u}_\lambda$ with N, and n_{Σ_τ} is the normal of Σ_τ. The left-hand side of (1.3) is called the N-*energy* of the Gaussian beam $\tilde{u}_\lambda$. The approximation in (1.3) can be made arbitrarily good and the time T arbitrarily large if we only take $\lambda > 0$ to be big enough. This characterisation of the energy allows then for a refinement of (1.2):[3]

> There exist (actual) solutions of the wave equation (1.1) whose N-energy is localised along a given null geodesic γ and behaves approximately like $-g(N, \dot{\gamma})\big|_{\mathrm{Im}(\gamma)\cap\Sigma_\tau}$ up to some finite, but arbitrarily large, time T. Here, $\dot{\gamma}$ is with respect to some affine parametrisation of γ. (1.4)

[1] See Theorem 2.1.

[2] We refer the reader to (1.8) in Section 1E for the definition of the stress–energy tensor.

[3] See Theorem 5.1.

It is worth emphasising that the need for an understanding of the temporal behaviour of the energy only arises for Gaussian beams on Lorentzian manifolds that *do not admit a globally timelike Killing vector field*[4] — otherwise there is a canonical energy which is conserved for solutions to the wave equation (1.1). Thus, for the majority of problems which so far found applications of Gaussian beams, for example the obstacle problem or the wave equation in time-independent inhomogeneous media, the question of the temporal behaviour of the energy did not arise (since it is trivial). However, understanding this behaviour on general Lorentzian manifolds is crucial for widening the application of Gaussian beams to problems arising, in particular, from general relativity.

In Part II, by applying (1.4), we derive some new results on the study of the wave equation on the familiar Schwarzschild, Reissner–Nordström, and Kerr black hole backgrounds (see [Hawking and Ellis 1973] for an introduction to these spacetimes):

(1) It is well-known folklore that the trapping[5] at the *photon sphere* in Reissner–Nordström and in Kerr necessarily leads to a "loss of derivative" in a local energy decay (LED) statement. We give a rigorous proof of this fact.

(2) We also show that the trapping at the *horizon* of an extremal Reissner–Nordström (and Kerr) black hole necessarily leads to a loss of derivative in an LED statement.

(3) When solving the wave equation (1.1) on the exterior of a Schwarzschild black hole backwards in time, the red-shift effect at the event horizon turns into a blue-shift: we construct solutions to the backwards problem whose energies grow exponentially for a finite, but arbitrarily long, time. This demonstrates the obstruction formed by the red-shift effect at the event horizon to scattering constructions from the future.

(4) Finally, we give a simple mathematical realisation of the heuristics for the blue-shift effect near the Cauchy horizon of (sub)extremal Reissner–Nordström and Kerr black holes: we construct a sequence of solutions to the wave equation whose initial energy is uniformly bounded whereas the energy near the Cauchy horizon goes to infinity.

Outline of the paper. We start by giving a short historical review of Gaussian beams in Section 1A. Thereafter we briefly explain how the notion of "energy" arises in the study of the wave equation and why it is important. We also discuss how the results we obtain allow us to disprove certain uniform statements about the temporal behaviour of the energy of waves. Section 1C elaborates on the wide applicability of the Gaussian beam approximation and explains its advantage over the geometric optics approximation. In the physics literature a similar "characterisation of the energy of high frequency waves" is folklore — we discuss its origin in Section 1D and put it into context with the work presented in this paper. Section 1E lays down the notation we use.

[4]One could add here "*uniformly*" timelike, meaning that the timelike Killing vector field does not "degenerate" when approaching the "boundary" of the manifold. Let us just state here that one can give precise meaning to "degenerating at the boundary".

[5]We do not intend to give a precise definition in this paper of what we mean by "trapping". However, loosely speaking "trapping" refers here to the presence of null geodesics that stay for all time in a compact region of "space".

Part I discusses the theory of Gaussian beams on Lorentzian manifolds. Sections 2 and 3 recall the construction of Gaussian beams and sketch the proof of Theorem 2.1, which basically says (1.2) and is more or less well known. In Section 4 we characterise the energy of a Gaussian beam, which is the main result of Part I. This result is then incorporated into Theorem 2.1, which yields Theorem 5.1 (or (1.4)). Moreover, Section 5 contains some general theorems which are tailored to the needs of many applications.

In Part II, we prove the above mentioned new results on the behaviour of waves on various black hole backgrounds. The important ideas are first introduced in Section 6 by the example of the Schwarzschild and Reissner–Nordström family, whose simple form of the metric allows for an uncomplicated presentation. Thereafter, in Section 7, we proceed to the Kerr family.

In the Appendix we give a sufficient criterion for the formation of caustics, i.e., a breakdown criterion for solutions of the eikonal equation, which shows the limitations of the "naive" geometric optics approximation.

1A. ***A brief historical review of Gaussian beams.*** The ansatz

$$u_\lambda = e^{i\lambda\phi}\left(a_0 + \frac{1}{\lambda}a_1 + \cdots + \frac{1}{\lambda^N}a_N\right) \tag{1.5}$$

for either a highly oscillatory approximate solution to some PDE or for a highly oscillatory approximate eigenfunction to some partial differential operator is known as the *geometric optics ansatz*. Here, $N \in \mathbb{N}$, ϕ is a real function (called the eikonal), the a_k are complex-valued functions, and λ is a positive parameter determining how quickly the function u_λ oscillates. In the widest sense, we understand under a Gaussian beam a function of the form (1.5) with a *complex*-valued eikonal ϕ that is real-valued along a bicharacteristic and has growing imaginary part off this bicharacteristic. This then leads to an exponential fall off in λ away from the bicharacteristic.

The use of a complex eikonal, although in a slightly different context, appears already in work of Keller [1956]. It was, however, only in the 1960s that the method of Gaussian beams was systematically applied and explored — mainly from a physics perspective. For more on these early developments we refer the reader to [Arnaud 1973, Chapter 4] and references therein. A general, mathematical theory of Gaussian beams, or what he called the complex WKB method, was developed by Maslov; see his book [1994] for an overview and also for references. Several of the later papers on Gaussian beams have their roots in this work.

The earliest application of the Gaussian beam method was to the construction of *quasimodes*; see, for example, [Ralston 1976]. Quasimodes approximately satisfy some type of Helmholtz equation, and thus they give rise to time-harmonic, approximate solutions to a wave equation. In this way quasimodes can be interpreted as standing waves. Later, various people used the Gaussian beam method for the construction of Gaussian *wave packets* (but also called "Gaussian beams") which form approximate solutions to a hyperbolic PDE.[6] Those wave packets, in contrast to quasimodes, are not stationary waves, but they move

[6]It is this sort of "Gaussian beam" that is the subject of this paper for the case of the wave equation on Lorentzian manifolds. More appropriately, one could name them "Gaussian wave packets" or "Gaussian pulses" to distinguish them from the standing waves — which are actually *beams*. However, we stick to the standard terminology.

through space, the trajectory in spacetime being a bicharacteristic of the partial differential operator. A detailed reference for this construction is [Ralston 1982], which goes back to 1977. Another presentation of this construction scheme was given by Babich and Ulin [1981].

Since then, there have been a lot of papers applying Gaussian beams to various problems.[7] For instance, in quantum mechanics Gaussian beams correspond to semiclassical approximate solutions to the Schrödinger equation and thus help understand the classical limit; or, in geophysics, one models seismic waves using the Gaussian beam approximation for solutions to a wave equation in an inhomogeneous (time-independent) medium.

1B. *Gaussian beams and the energy method.*

1B1. *The energy method as a versatile method for studying the wave equation.* The study of the wave equation on various geometries has a long history in mathematics and physics. A very successful and widely applicable method for obtaining quantitative results on the long-time behaviour of waves is the energy method. It was pioneered by Morawetz [1961; 1962], where she proved pointwise decay results in the context of the obstacle problem. In [Morawetz 1968] she established what is now known as *integrated local energy decay* (ILED) for solutions of the Klein–Gordon equation (and thus inferring decay). In the past ten years, her methods were adapted and extended by many people in order to prove boundedness and decay of waves on various (black hole) spacetimes — a study which is mainly motivated by the black hole stability conjecture (see the introduction of [Dafermos and Rodnianski 2013]). A small selection of examples is [Klainerman 1985; Dafermos and Rodnianski 2009; 2010a; 2011a; 2011b; Andersson and Blue 2009; Tataru and Tohaneanu 2011; Luk 2010; Schlue 2013; Aretakis 2011a; Holzegel and Smulevici 2013; Civin 2014; Dyatlov 2011].

The philosophy of the energy method is first to derive estimates on a suitable energy (and higher-order energies)[8] and then to establish pointwise estimates using Sobolev embeddings. Thus, given a spacetime on which one intends to study the wave equation using the energy method, one first has to set up such a suitable energy (and higher-order energies — but in this paper we focus on the first-order energy). A general procedure is to construct an energy from a foliation of the spacetime by spacelike slices Σ_τ together with a timelike vector field N; see (1.9) in Section 1E. We refrain from discussing here what choices of foliation and timelike vector field lead to a "suitable" notion of energy.[9] Let us just mention here that, in the presence of a globally timelike Killing vector field T, one obtains a particularly well-behaved energy by choosing $N = T$ and a foliation that is invariant under the flow of T.[10] We invite the reader to convince him- or herself that the familiar notions of energy for the wave equation on the Minkowski spacetime or in time-independent inhomogeneous media arise as special cases of this more general scheme.

[7] We refer the reader to [Maslov 1994] for a list of references.

[8] A *first-order* energy controls the *first* derivatives of the wave and is referred to in the following just as "energy". *Higher-order* energies control *higher* derivatives of the wave. A special case of the energy method is the so-called *vector field method*. Higher-order energies arise there naturally by commutation with suitable vector fields; see [Klainerman 1985].

[9] However, see Section II for some examples and footnote 27 on page 1400 for some further comments.

[10] Such a choice corresponds to what we called in the introduction a "canonical energy".

1B2. *Gaussian beams as obstructions to certain uniform behaviour of the energy of waves.* The approximation with Gaussian beams allows us to construct solutions to the wave equation whose energy is localised for an arbitrarily long, but finite, time along a null geodesic. Such solutions naturally form an obstruction to certain uniform statements about the temporal behaviour of the energy of waves. A classical example is the case in which one has a null geodesic that does not leave a compact region in "space" and which has constant energy.[11] Such null geodesics form obstructions to certain formulations of local energy decay being true.[12] However, it is very important to be aware of the fact that, in general, none of the solutions from (1.4) has localised energy for *all* time. Thus, in order to contradict, for instance, an LED statement, it is in general inevitable to resort to a *sequence* of solutions of the form (1.4) which exhibit the contradictory behaviour in the limit. For this scheme to work, however, it is clearly crucial that the LED statement in question is *uniform* with respect to some energy which is left constant by the sequence of Gaussian beam solutions. Note here that (1.4) states in particular that the time T up to which one has good control over the wave can be made arbitrarily large without changing the initial energy! Higher-order initial energies, however, will blow up when T is taken bigger and bigger. In this paper we restrict our consideration to disproving statements that are uniform with respect to the first-order energy. In Sections 6A, 6F and 7A, we demonstrate this important application of Gaussian beams: we show that certain (I)LED statements derived by various people in the presence of "trapping" are sharp in the sense that *some* loss of derivative is necessary (however, one does not necessarily need to lose a whole derivative; see the discussion at the end of Section 6A).

We conclude this section with the remark that in the presence of a globally timelike Killing vector field one can already infer such obstructions from (1.2), since the (canonical) energy of solutions to the wave equation is then constant. In this way, one can easily infer from (1.2) alone that an LED statement in Schwarzschild has to lose differentiability due to the trapping at the photon sphere. But already for trapping in Kerr one needs to know how the "trapped" energy of the solutions referred to in (1.2) behaves in order to infer the analogous result. This knowledge is provided by (1.3) and/or (1.4).

1C. ***Gaussian beams are parsimonious.*** The approximation by Gaussian beams can be carried out on a Lorentzian manifold (M, g) under minimal assumptions:

(1) One needs a well-posed initial value problem. This is ensured by requiring that (M, g) is globally hyperbolic.[13] However, one can also replace the well-posed initial value problem by a well-posed initial–boundary value problem — and one can obtain, with small changes and some additional work in the proof, qualitatively identical results.

(2) Having fixed an N-energy to work with, one has to have an energy estimate of the form (2.8) at one's disposal, which is guaranteed by the condition (2.3). The estimate (2.8) allows us to infer that the approximation by the Gaussian beam is *global* in space. It is only under this condition that it is justified to

[11] We refer to the right-hand side of (1.3) as the N-energy of the null geodesic.

[12] A classic regarding such a result is by Ralston [1969]. However, he does not use the Gaussian beam approximation in this work, but the geometric optics approximation.

[13] The assumption of global hyperbolicity has another simplifying, but not essential, feature; see the discussion after Definition 3.13.

say in (1.2) and (1.4) that the energy of the actual solution is *localised* along a null geodesic.[14] However, as we show in Remark 2.9, one *always* has a *local* approximation, which is, together with the geometric characterisation of the energy, sufficient for obtaining control of the wave in a small neighbourhood of the underlying null geodesic *regardless of condition* (2.3). This then allows us to establish, for example, the very general Theorem 5.5, which only requires global hyperbolicity (or some other form of well-posedness for the wave equation; see (1)).

In particular, the method of Gaussian beams is not in need of any special structure on the Lorentzian manifold like Killing vector fields (as, for example, needed for the mode analysis or for the construction of quasimodes).

We would also like to emphasise here that in order to apply (1.4) one only needs to understand the behaviour of the null geodesics of the underlying Lorentzian manifold! This knowledge is often in reach and thus Gaussian beams provide in many cases an easy and feasible way for obtaining control of highly oscillatory solutions to the wave equation. In this sense the theory presented in Part I forms a good "black box result" which can be applied to various different problems.

We conclude this section with a brief comparison of the Gaussian beam approximation with the geometric optics approximation: Let us call the geometric optics approximation, which considers approximate solutions of the form (1.5), the "naive" geometric optics approximation. Although it applies under the same general conditions as the Gaussian beam approximation, in general the time T up to which one has good control over the solution cannot be chosen arbitrarily large, since the approximate solution breaks down at caustics. In the Appendix we show that caustics necessarily form along null geodesics that possess conjugate points. A prominent example of such null geodesics are the trapped null geodesics at the photon sphere in the Schwarzschild spacetime (see Section 6A for the proof that these null geodesics have conjugate points). However, the formation of caustics is not a serious limitation of the geometric optics approximation, since one can extend the approximate solution through the caustics, making use of Maslov's canonical operator. The approximate solution obtained in this way is, however, no longer of the simple form (1.5). The advantage of the Gaussian beam approximation is that the simple ansatz (1.5) does not break down at caustics; it yields an approximation up to all finite times T.

1D. ***"High-frequency" waves in the physics literature.*** In physics, the notion of a *local observer's energy* arose with the emergence of Einstein's theory of relativity: Suppose an observer travels along a timelike curve $\sigma : I \to M$ with unit velocity $\dot{\sigma}$. Then, with respect to a Lorentz frame of his, he measures the local energy density of a wave u to be $\mathbb{T}(u)(\dot{\sigma}, \dot{\sigma})$, where $\mathbb{T}(u)$ is the stress–energy tensor of the wave u; see (1.8) in Section 1E. By considering the 3-parameter family of observers whose velocity vector field is given by the normal n_{Σ_τ} to a foliation of M by spacelike slices Σ_τ, the physical definition of energy is contained in the mathematical one (which is given by (1.9)).

[14]That one needs condition (2.3) for ensuring that the energy is indeed localised is in fact another minor novelty in the study of Gaussian beams on *general* Lorentzian manifolds (note that, in the case of N being a Killing vector field, condition (2.3) is trivially satisfied). For an example for a violation of condition (2.3) we refer to the discussion after (6.8) on page 1406.

The prevalent description of highly oscillatory (or "high-frequency") waves in the physics literature is that the waves (or "photons") propagate along null geodesics γ and each of these rays (or photons) carries an energy–momentum 4-vector $\dot{\gamma}$, where the dot is with respect to some affine parametrisation. In the high-frequency limit, the number of photons is preserved. Thus, the energy of the wave, as measured by a local observer with world line σ, is determined by the energy component $-g(\dot{\gamma}, \dot{\sigma})$ of the momentum 4-vector $\dot{\gamma}$. By considering a highly oscillatory wave that "gives rise to just one photon", one recovers the characterisation of the energy of a Gaussian beam, (1.3), given in this paper.

In the physics literature (see, for example, the classic [Misner et al. 1973, Chapter 22.5]), this description is justified using the naive geometric optics approximation. Here, it suffices to take $N = 0$ in (1.5); one then considers approximate solutions to the wave equation of the form $u_\lambda = a \cdot e^{i\lambda\phi}$, where a and ϕ satisfy

$$d\phi \cdot d\phi = 0 \quad \text{and} \quad 2\,\mathrm{grad}\,\phi(a) + \Box\phi \cdot a = 0. \tag{1.6}$$

The conservation law

$$\mathrm{div}(a^2\,\mathrm{grad}\,\phi) = 0, \tag{1.7}$$

which can be easily inferred from the second equation in (1.6), is interpreted as the conservation of the number-flux vector $S = a^2\,\mathrm{grad}\,\phi$ of the photons. The leading component in λ of the renormalised[15] stress–energy tensor $T(u_\lambda)$ of the wave $u_\lambda = a \cdot e^{i\lambda\phi}$ in the geometric optics limit is then given by

$$\mathbb{T}(u_\lambda) = \mathrm{grad}\,\phi \otimes S,$$

from which it then follows that each photon carries a 4-momentum $\mathrm{grad}\,\phi = \dot{\gamma}$.

In particular, making use of the conservation law (1.7), it is not difficult[16] to prove a geometric characterisation of the energy of waves in the naive geometric optics limit analogous to the one we prove in this paper for Gaussian beams. However, as we have mentioned in the previous section, the naive geometric optics approximation has the undesirable feature that it breaks down at caustics.

The characterisation of the energy of Gaussian beams is more difficult, since (1.7) is replaced only by an *approximate* conservation law.[17] Moreover, it provides a rigorous justification of the temporal behaviour of the local observer's energy of photons, *which also applies to photons along whose trajectory caustics would form.*

1E. *Notation.* Given a Lorentzian manifold (M, g), we denote the canonical isomorphisms induced by the metric g between the tangent and cotangent space by $\sharp : T^*_x M \to T_x M$ and $\flat : T_x M \to T^*_x M$, where $x \in M$ and, for $\alpha \in T^*_x M$ and $X \in T_x M$, the isomorphisms $\sharp$ and $\flat$ are given by $\alpha^\sharp := g^{-1}(\alpha, \cdot)$ and $X^\flat := g(X, \cdot)$. Here g^{-1} denotes the inverse of the metric g. Moreover, we denote with $\cdot$ the inner product of two vectors as well as the inner product of two covectors, i.e., for $\alpha, \beta \in T^*_x M$ we write $\alpha \cdot \beta := g^{-1}(\alpha, \beta)$, and for $X, Y \in T_x M$ we write $X \cdot Y := g(X, Y)$. We also introduce the notation $\mathrm{grad}\, f := (df)^\sharp$ for the gradient of a function $f \in C^\infty(M, \mathbb{R})$. The Levi-Civita connection on the

[15] Divided by λ^2.

[16] Although, to the best of our knowledge, it is nowhere done explicitly.

[17] See the discussion below (4.6) in Section 4.

Lorentzian manifold (M, g) is denoted by ∇, and we write $\operatorname{div} Z := \nabla_\mu Z^\mu$ for the divergence of a smooth vector field Z on M. Furthermore, we define the *wave operator* $\Box_g$ by

$$\Box_g u := \nabla^\mu \nabla_\mu u.$$

From here on we will, however, omit the index g on $\Box_g$, since it is clear from the context which Lorentzian metric is referred to.

Whenever we are given a time-oriented Lorentzian manifold (M, g) that is (partly) foliated by spacelike slices $\{\Sigma_\tau\}_{\tau\in[0,\tau^*)}$, $0<\tau^*\leq\infty$, we denote the future-directed unit normal to the slice Σ_τ by n_{Σ_τ}. Moreover, the induced Riemannian metric on Σ_τ is then denoted by $\bar{g}_\tau$ and we set $R_{[0,T]} := \bigcup_{0\leq\tau\leq T} \Sigma_\tau$.

For $u \in C^\infty(M, \mathbb{C})$ we define the *stress–energy tensor* $\mathbb{T}(u)$ by

$$\mathbb{T}(u) := \tfrac{1}{2}\overline{du} \otimes du + \tfrac{1}{2} du \otimes \overline{du} - \tfrac{1}{2} g(\,\cdot\,,\cdot\,) g^{-1}(du, \overline{du}). \tag{1.8}$$

Given in addition a vector field N, we define the current $J^N(u)$ by

$$J^N(u) := [\mathbb{T}(u)(N, \cdot\,)]^\sharp.$$

Finally, if N is future-directed timelike, we call

$$E^N_\tau(u) := \int_{\Sigma_\tau} J^N(u) \cdot n_{\Sigma_\tau} \operatorname{vol}_{\bar{g}_\tau} \tag{1.9}$$

the *N-energy* of u at time τ, where $\operatorname{vol}_{\bar{g}_\tau}$ denotes the volume element corresponding to the metric $\bar{g}_\tau$.[18] If $A \subseteq \Sigma_\tau$, then $E^N_{\tau,A}(u)$ denotes the N-energy of u at time τ in the volume A, i.e., the integration in (1.9) is only over A.

The notion (1.9) of the N-energy of a function u is especially helpful whenever we have an adequate knowledge of $\Box u$, since one can then infer detailed information about the behaviour of the N-energy (see the energy estimate (2.8) in the next section), and thus also about the behaviour of u itself. Hence, the stress–energy tensor (1.8) together with the notion of the N-energy is particularly useful for solutions u of the *wave equation*

$$\Box u = 0. \tag{1.10}$$

For more on the stress–energy tensor and the notion of energy, we refer the reader to [Taylor 2011, Chapters 2.7 and 2.8].

Given a Lorentzian manifold (M, g) and $A \subseteq M$, we denote with $J^+(A)$ the causal future of A, namely, all the points $x \in M$ such that there exists a future-directed causal curve starting at some point of A and ending at x. The causal past of A, $J^-(A)$, is defined analogously.[19] Finally, C and c will always denote positive constants.

For simplicity of notation we restrict our considerations to 3+1-dimensional Lorentzian manifolds (M, g). However, all results extend in an obvious way to dimensions $n+1$, $n \geq 1$. Moreover, all given

[18] See also [Choquet-Bruhat 2009, Appendix III, Sections 2.3 and 2.4] (in particular Definition (2.27)) for a detailed discussion of the notion of N-energy.

[19] See also Chapter 14 in [O'Neill 1983].

manifolds, functions and tensor fields are assumed to be smooth, although this is only for convenience and clearly not necessary.

Part I. The theory of Gaussian beams on Lorentzian manifolds

2. Solutions of the wave equation with localised energy

This section and the next are devoted to a sketch of the proof of Theorem 2.1, which summarises the state of the art knowledge concerning the construction of solutions with localised energy using the approximation by Gaussian beams.

Theorem 2.1. *Let (M, g) be a time-oriented, globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$, where Σ_0 is a Cauchy hypersurface.*[20] *Furthermore, let γ be a null geodesic that intersects Σ_0 and N a timelike, future-directed vector field.*

For any neighbourhood $\mathcal{N}$ of γ, any $T > 0$ with $\Sigma_T \cap \mathrm{Im}(\gamma) \neq \varnothing$ (see Figure 1), and any $\mu > 0$, there exists a solution $v \in C^\infty(M, \mathbb{C})$ of the wave equation (1.10) with $E^N_0(v) = 1$ and $\tilde{u} \in C^\infty(M, \mathbb{C})$ with $\mathrm{supp}(\tilde{u}) \subseteq \mathcal{N}$ *such that*

$$E^N_\tau(v - \tilde{u}) < \mu \quad \textit{for all } 0 \leq \tau \leq T \tag{2.2}$$

provided that we have, on $R_{[0,T]} \cap J^+(\mathcal{N} \cap \Sigma_0)$,

$$\frac{1}{|dt(n_{\Sigma_\tau})|} + |g(N, n_{\Sigma_\tau})| \leq C < \infty \quad \text{and} \quad 0 < c \leq |g(N, N)|,$$

$$|\nabla N(n_{\Sigma_\tau}, n_{\Sigma_\tau})| + \sum_{i=1}^{3} |\nabla N(n_{\Sigma_\tau}, e_i)| + \sum_{i,j=1}^{3} |\nabla N(e_i, e_j)| \leq C < \infty, \tag{2.3}$$

where c and C are positive constants and $\{n_{\Sigma_\tau}, e_1, e_2, e_3\}$ is an orthonormal frame.

Note that (2.2) together with $\mathrm{supp}(\tilde{u}) \subseteq \mathcal{N}$ make rigorous the statement that the solution v hardly disperses up to time T. The energy of the solution v stays localised for finite time.

Proof. The function $\tilde{u}$ in the theorem is the Gaussian beam, the *approximate* solution to the wave equation (1.10) which we need to construct. Recall that a Gaussian beam $u_\lambda \in C^\infty(M, \mathbb{C})$ is of the form

$$u_\lambda(x) = a_{\mathcal{N}}(x) e^{i\lambda\phi(x)}, \tag{2.4}$$

where $\lambda > 0$ is a parameter that determines how quickly the Gaussian beam oscillates, and $a_{\mathcal{N}}$ and ϕ are smooth, complex-valued functions on M that do *not* depend on λ. However, $a_{\mathcal{N}}$ depends on the neighbourhood $\mathcal{N}$ of the null geodesic γ. In Section 3 we outline how one constructs the functions $a_{\mathcal{N}}$ and ϕ in such a way that u_λ satisfies the following three conditions: The *first condition* is

$$\|\Box u_\lambda\|_{L^2(R_{[0,T]})} \leq C(T), \tag{2.5}$$

[20] Bernal and Sánchez [2005] showed that every globally hyperbolic Lorentzian manifold admits a smooth time function.

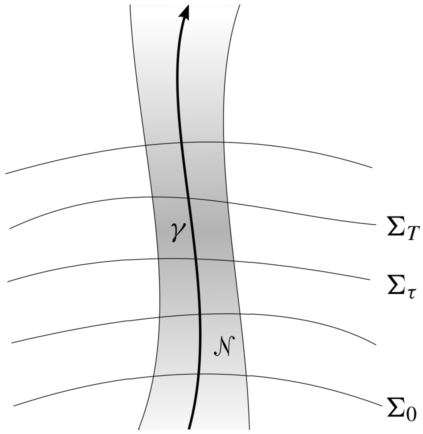

Figure 1. The setting of Theorem 2.1.

where the constant $C(T)$ depends on $a_{\mathcal{N}}$, ϕ and T, but not on λ. The *second condition* is

$$E_0^N(u_\lambda) \to \infty \quad \text{for } \lambda \to \infty, \tag{2.6}$$

where N is the timelike vector field from Theorem 2.1. Finally, the *third condition* is

$$u_\lambda \text{ is supported in } \mathcal{N}. \tag{2.7}$$

Assuming for now that we have already found functions $a_{\mathcal{N}}$ and ϕ such that the conditions (2.5), (2.6) and (2.7) are satisfied, we finish the proof of Theorem 2.1. In order to normalise the initial energy of the approximate solutions u_λ, we define

$$\tilde{u}_\lambda := \frac{u_\lambda}{\sqrt{E_0^N(u_\lambda)}},$$

which, moreover, yields

$$\|\Box \tilde{u}_\lambda\|_{L^2(R_{[0,T]})} \to 0 \quad \text{for } \lambda \to \infty.$$

This says that as the Gaussian beam becomes more and more oscillatory (i.e., for bigger and bigger λ), the closer it comes to being a proper solution to the wave equation.

We now define the actual solution v_λ of the wave equation — the one that is being approximated by the $\tilde{u}_\lambda$ — to be the solution of the following initial value problem:

$$\begin{aligned} \Box v &= 0, \\ v\big|_{\Sigma_0} &= \tilde{u}_\lambda\big|_{\Sigma_0}, \\ n_{\Sigma_0} v\big|_{\Sigma_0} &= n_{\Sigma_0}\tilde{u}_\lambda\big|_{\Sigma_0}. \end{aligned}$$

Here, we make use of the fact that the Lorentzian manifold (M, g) is globally hyperbolic and thus allows for a well-posed initial value problem for the wave equation. Moreover, the condition (2.3) ensures that we have an energy estimate of the form

$$\int_{\Sigma_\tau} J^N(u)\cdot n_{\Sigma_\tau}\,\mathrm{vol}_{\bar{g}_\tau} \le C(T,N,\{\Sigma_\tau\})\Big(\int_{\Sigma_0} J^N(u)\cdot n_{\Sigma_0}\,\mathrm{vol}_{\bar{g}_0} + \|\Box u\|^2_{L^2(R_{[0,T]})}\Big) \quad \text{for all } 0\le\tau\le T \tag{2.8}$$

at our disposal (see for example [Taylor 2011, Chapter 2.8]). Thus, we obtain

$$E^N_\tau(v_\lambda - \tilde{u}_\lambda) \le C(T,N,\Sigma_\tau)\cdot \|\Box\tilde{u}_\lambda\|^2_{L^2(R_{[0,T]})} \quad \text{for all } 0\le\tau\le T,$$

which goes to zero for $\lambda\to\infty$. Given now $\mu>0$, it suffices to choose $\lambda_0>0$ big enough and to set $\tilde{u} := \tilde{u}_{\lambda_0}$ and $v := v_{\lambda_0}$, which then finishes the proof under the assumption of the conditions (2.5), (2.6) and (2.7). □

We end this section with a couple of remarks about Theorem 2.1:

Remark 2.9. As already mentioned, the condition (2.3) ensures that we have the energy estimate (2.8). It is automatically satisfied if the region under consideration, $R_{[0,T]}\cap J^+(\mathcal{N}\cap\Sigma_0)$, is relatively compact, which will be the case in many concrete applications.

Moreover, by choosing $\mathcal{N}$ a bit smaller if necessary, we can always arrange that $\Sigma_T\cap\mathcal{N}$ is relatively compact and that $\mathcal{N}\cap R_{[0,T]}\subseteq J^-(\Sigma_T\cap\mathcal{N})$. Doing, then, the energy estimate in the relatively compact region $J^-(\Sigma_T\cap\mathcal{N})\cap J^+(\Sigma_0)$, we obtain

$$E^N_{\tau,\mathcal{N}\cap\Sigma_\tau}(v-\tilde{u}) < \mu \quad \text{for all } 0\le\tau\le T \tag{2.10}$$

independently of (2.3). Of course, the information given by (2.10) is not interesting here, since Theorem 2.1 does not provide more information about $\tilde{u}$ than its region of support. However, in Section 4 we will derive more information about the approximate solution $\tilde{u}$ and then (2.10) will tell us about the temporal behaviour of the localised energy of v; see Theorem 5.1.

Remark 2.11. By taking the real or the imaginary part of $\tilde{u}_\lambda$ and v_λ it is clear that we can choose $\tilde{u}$ and v in Theorem 2.1 to be real valued.

3. The construction of Gaussian beams

Before we sketch the construction of Gaussian beams, let us mention that other (and complete) presentations of this subject can be found, for example, in [Babich and Buldyrev 2009] or [Ralston 1982]. The latter reference also includes the construction of Gaussian beams for more general hyperbolic PDEs.

Given now a neighbourhood $\mathcal{N}$ of a null geodesic γ, we need to construct functions $a_{\mathcal{N}},\phi\in C^\infty(M,\mathbb{C})$ such that the approximate solution $u_\lambda = a_{\mathcal{N}}\cdot e^{i\lambda\phi}$ satisfies the conditions (2.5), (2.6) and (2.7). This will then finish the proof of Theorem 2.1. We compute

$$\Box u_\lambda = -\lambda^2(d\phi\cdot d\phi)a_{\mathcal{N}}e^{i\lambda\phi} + i\lambda\Box\phi\cdot a_{\mathcal{N}}e^{i\lambda\phi} + 2i\lambda\ \mathrm{grad}\,\phi(a_{\mathcal{N}})\cdot e^{i\lambda\phi} + \Box a_{\mathcal{N}}\cdot e^{i\lambda\phi}. \tag{3.1}$$

Demanding $d\phi\cdot d\phi = 0$ (the eikonal equation) and $2\,\mathrm{grad}\,\phi(a_{\mathcal{N}}) + \Box\phi\cdot a_{\mathcal{N}} = 0$ would lead us to the naive geometric optics approximation (see (1.6)), whose major drawback is that in general the solution ϕ of the eikonal equation breaks down at some point along γ due to the formation of caustics. The method of Gaussian beams takes a slightly different approach. We only require an *approximate* solution $\phi \in C^\infty(M, \mathbb{C})$ of the eikonal equation in the sense that

$$d\phi \cdot d\phi \;\text{ vanishes on } \gamma \text{ to high order.}[21]$$

Moreover, we demand that

$$\phi\big|_\gamma \;\text{ and }\; d\phi\big|_\gamma \;\text{ are real valued,} \tag{3.2}$$

$$\mathfrak{Im}\big(\nabla\nabla\phi\big|_\gamma\big) \;\text{ is positive definite on a 3-dimensional subspace transversal to }\; \dot{\gamma}, \tag{3.3}$$

where $\mathfrak{Im}\big(\nabla\nabla\phi\big|_x\big)$, $x \in M$, denotes the imaginary part of the bilinear map $\nabla\nabla\phi\big|_x : T_xM \times T_xM \to \mathbb{C}$. Let us assume for a moment that (3.2) and (3.3) hold. Taking slice coordinates for γ, that is, a coordinate chart (U, φ), $\varphi : U \subseteq M \to \mathbb{R}^4$, such that $\varphi(\mathrm{Im}(\gamma) \cap U) = \{x_1 = x_2 = x_3 = 0\}$, we obtain

$$\mathfrak{Im}(\phi)(x) \geq c \cdot (x_1^2 + x_2^2 + x_3^2), \tag{3.4}$$

at least if we restrict ϕ to a small enough neighbourhood of γ. Note that such slice coordinates exist, since the global hyperbolicity of (M, g) implies that γ is an embedded submanifold of M. This is easily seen by appealing to the strong causality condition.[22] Let us now denote the real part of ϕ by ϕ_1 and the imaginary part by ϕ_2. We then have

$$u_\lambda = a_{\mathcal{N}} \cdot e^{i\lambda\phi_1} \cdot e^{-\lambda\phi_2}.$$

We see that the last factor imposes the shape of a Gaussian on u_λ, centred around γ — this explains the name. Moreover, for λ large this Gaussian will become more and more narrow, i.e., less and less weight is given to the values of $a_{\mathcal{N}}$ away from γ.

We rewrite (3.1) as

$$\Box u_\lambda = -\lambda^2 \underbrace{(d\phi \cdot d\phi)}\cdot a_{\mathcal{N}} e^{i\lambda\phi_1} \cdot e^{-\lambda\phi_2} + i\lambda\underbrace{(2\,\mathrm{grad}\,\phi\,(a_{\mathcal{N}}) + \Box\phi \cdot a_{\mathcal{N}})}\cdot e^{i\lambda\phi_1} \cdot e^{-\lambda\phi_2} + \Box a_{\mathcal{N}} \cdot e^{i\lambda\phi_1} \cdot e^{-\lambda\phi_2}. \tag{3.5}$$

Intuitively, if we can arrange for the underbraced terms to vanish on γ to some order and we choose large λ, then we will pick up only very small contributions. The next lemma makes this rigorous:

Lemma 3.6. *Let $f \in C_0^\infty([0, T] \times \mathbb{R}^3, \mathbb{C})$ vanish along $\{x_1 = x_2 = x_3 = 0\}$ to order S, that is, all partial derivatives up to and including the order S of f vanish along $\{x_1 = x_2 = x_3 = 0\}$, and let $c > 0$ be a constant. We then have*

(i) $$\int_{[0,T]\times\mathbb{R}^3} |f(x)|^2 e^{-\lambda\cdot c(x_1^2+x_2^2+x_3^2)}\, dx \leq C\lambda^{-(S+1)-3/2}$$

and

(ii) $$\int_{[0,T]\times\mathbb{R}^3} |f(x)| e^{-\lambda\cdot c(x_1^2+x_2^2+x_3^2)}\, dx \leq C\lambda^{-(S+1)/2-3/2},$$

where C depends on f (and on T).

[21] The exact order to which we require $d\phi \cdot d\phi$ to vanish on γ will be determined later.

[22] See, for example, [O'Neill 1983, Chapter 14] for more on the strong causality condition.

Proof. We prove (i) here, since it is used in the following. The formulation (ii) of Lemma 3.6 is appealed to in the proof of Theorem 4.1 in Section 4 — the proof is analogous.

Introduce stretched coordinates $y_0 := x_0$, $y_i := \sqrt{\lambda} x_i$ for $i = 1, 2, 3$. Since f vanishes along the x_0 axis to order S and has compact support, we get $|f(x)| \leq C \cdot |\underline{x}|^{S+1}$ for all $x = (x_0, \underline{x}) \in [0, T] \times \mathbb{R}^3$; thus

$$\left| f\left(y_0, \frac{\underline{y}}{\sqrt{\lambda}}\right)\right| \leq C \cdot \frac{|\underline{y}|^{S+1}}{\lambda^{(S+1)/2}}.$$

This yields

$$\int_{[0,T]\times\mathbb{R}^3} |f(x)|^2 e^{-\lambda \cdot c|\underline{x}|^2}\, dx \leq \int_{[0,T]\times\mathbb{R}^3} C \cdot |\underline{y}|^{2(S+1)} e^{-c|\underline{y}|^2}\, dy \cdot \lambda^{-(S+1)-3/2}. \tag{3.7}$$

This concludes the proof. □

We summarise the approach taken by the Gaussian beam approximation in the following:

Lemma 3.8. *Within the setting of Theorem 2.1, assume we are given $a, \phi \in C^\infty(M, \mathbb{C})$ which satisfy (3.2) and (3.3). Moreover, assume*

$$d\phi \cdot d\phi \qquad \text{vanishes to second order along } \gamma, \tag{3.9}$$

$$2 \operatorname{grad} \phi(a) + \Box\phi \cdot a \qquad \text{vanishes to zeroth order along } \gamma, \tag{3.10}$$

$$a(\operatorname{Im}(\gamma) \cap \Sigma_0) \neq 0 \quad \text{and} \quad d\phi\, (\operatorname{Im}(\gamma) \cap \Sigma_0) \neq 0. \tag{3.11}$$

Given a neighbourhood $\mathcal{N}$ of γ, we can then multiply a by a suitable bump function $\chi_{\mathcal{N}}$, which is equal to one in a neighbourhood of γ and satisfies $\operatorname{supp}(\chi_{\mathcal{N}}) \subseteq \mathcal{N}$, *such that*

$$u_\lambda = u_{\lambda,\mathcal{N}} = a_{\mathcal{N}} e^{i\lambda\phi}$$

satisfies (2.5), (2.6) and (2.7), where $a_{\mathcal{N}} := a \cdot \chi_{\mathcal{N}}$.

Proof. Cover γ by slice coordinate patches and let $\tilde{\chi}$ be a bump function which meets the following three requirements:

(i) $\tilde{\chi}$ is equal to one in a neighbourhood of γ.

(ii) (3.4) is satisfied for all $x \in \operatorname{supp}(\tilde{\chi})$.

(iii) $R_{[0,T]} \cap \operatorname{supp}(\tilde{\chi})$ is relatively compact in M for all $T > 0$ with $\Sigma_T \cap \operatorname{Im}(\gamma) \neq \varnothing$.

Pick now a second bump function $\tilde{\chi}_{\mathcal{N}}$ which is again equal to one in a neighbourhood of γ and is supported in $\mathcal{N}$. We then define $\chi_{\mathcal{N}} := \tilde{\chi} \cdot \tilde{\chi}_{\mathcal{N}}$. Clearly, (2.7) is satisfied.

In order to see that (2.5) holds, note that the conditions (3.2), (3.3), (3.9) and (3.10) are still satisfied by the pair $(a_{\mathcal{N}}, \phi)$. Moreover note that, due to condition (iii), the integrand is supported in a compact region for each $T > 0$ with $\Sigma_T \cap \operatorname{Im}(\gamma) \neq \varnothing$. Thus, the spacetime volume of this region is finite. We thus obtain (2.5) from (3.5) and Lemma 3.6.

Finally, we have

$$E_0^N(u_\lambda) \geq C \cdot (\lambda^{1/2} - 1).$$

This follows since the highest-order term in λ in $E_0^N(u_\lambda)$ is

$$\lambda^2 \cdot \int_{\Sigma_0} |a_{\mathcal{N}}|^2 N\phi_1 \cdot n_{\Sigma_0}\phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_0},$$

and the same scaling argument used in the proof of Lemma 3.6 shows that the term $e^{-2\lambda\phi_2}$ leads to a $\lambda^{-3/2}$ damping — and only to a $\lambda^{-3/2}$ damping due to condition (3.11) (together with (3.9) and (3.2)). Thus, (2.6) is satisfied as well and the lemma is proved. □

For the actual construction of the functions ϕ and a such that (3.2), (3.3), (3.9), (3.10), and (3.11) are satisfied, we refer the reader, for example, to [Ralston 1982]. We content ourselves here with pointing out that the above conditions on the functions ϕ and a are actually only conditions on the first, second, and third derivatives of ϕ along γ and on the first derivative of a and the value of a itself along γ. Making the choice

$$d\phi(s) := \dot{\gamma}^\flat(s) \tag{3.12}$$

along γ, where s is an affine parameter for γ, the condition (3.9) turns into a quadratic ODE for the second derivatives of ϕ along γ, while the condition (3.10) turns into a linear ODE for a along γ. The important step is to show that one can find a *global* solution for the first ODE, which, moreover, also satisfies (3.3).

We conclude this section by making the following definition for future reference:

Definition 3.13. Let (M, g) be a time-oriented, globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$. Furthermore, let $\gamma : [0, S) \to M$ be an affinely parametrised future-directed null geodesic with $\gamma(0) \in \Sigma_0$, where $0 < S \leq \infty$, and let N be a timelike, future-directed vector field.

Given functions $a, \phi \in C^\infty(M, \mathbb{C})$ that satisfy (3.2), (3.3), (3.9), (3.10), $a(\operatorname{Im}(\gamma) \cap \Sigma_0) \neq 0$ and (3.12), we call the function

$$u_{\lambda,\mathcal{N}} = a_{\mathcal{N}} e^{i\lambda\phi}$$

a *Gaussian beam along γ with structure functions a and ϕ and with parameters λ and $\mathcal{N}$*. Here, $a_{\mathcal{N}} = a \cdot \chi_{\mathcal{N}} = a \cdot \tilde{\chi} \cdot \tilde{\chi}_{\mathcal{N}}$ with $\tilde{\chi}$ and $\tilde{\chi}_{\mathcal{N}}$ as in the proof of Lemma 3.8. Moreover, we call the function

$$\tilde{u}_{\lambda,\mathcal{N}} = \frac{u_{\lambda,\mathcal{N}}}{\sqrt{E_0^N(u_{\lambda,\mathcal{N}})}} \cdot \sqrt{E}$$

a *Gaussian beam along γ with structure functions a and ϕ, parameters λ and $\mathcal{N}$, and initial N-energy E*, where E is a strictly positive real number. Let us emphasise that, when we say "a Gaussian beam along γ", γ encodes here not only the image of γ, but also the affine parametrisation.

We end this section with the remark that, for the sole *construction* of the Gaussian beams, the assumption of the global hyperbolicity of (M, g) can be replaced by the assumption that the null geodesic $\gamma : \mathbb{R} \supseteq I \to M$ is a smooth embedding, in particular $\gamma(I)$ is an embedded submanifold. Moreover, note that, if $\gamma : \mathbb{R} \supseteq I \to M$ is a smooth injective immersion and if $[a, b] \subseteq I$ with $a, b \in \mathbb{R}$, then $\gamma|_{(a,b)} : (a, b) \to M$ is a smooth embedding. Thus the construction of a Gaussian beam is always possible for null geodesics with no self-intersections on general Lorentzian manifolds — at least up to some finite affine time in the domain of γ.

4. Geometric characterisation of the energy of Gaussian beams

In this section we characterise the energy of a Gaussian beam in terms of the energy of the underlying null geodesic. The following theorem is the main result of Part I:

Theorem 4.1. *Let (M, g) be a time-oriented, globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$. Moreover, let N be a timelike future-directed vector field and $\gamma : [0, S) \to M$ an affinely parametrised future-directed null geodesic with $\gamma(0) \in \Sigma_0$, where $0 < S \le \infty$.*

For any $T > 0$ with $\mathrm{Im}(\gamma) \cap \Sigma_T \neq \varnothing$ and any $\mu > 0$, there exists a $\lambda_0 > 0$ such that any Gaussian beam $\tilde{u}_{\lambda,\mathcal{N}}$ along γ with structure functions a and ϕ, parameters $\lambda \ge \lambda_0$ and $\mathcal{N}$, and initial N-energy equal to $-g(N, \dot{\gamma})\big|_{\gamma(0)}$, satisfies

$$\left|E^N_\tau(\tilde{u}_{\lambda,\mathcal{N}}) - \left(-g(N, \dot{\gamma})\big|_{\mathrm{Im}(\gamma)\cap\Sigma_\tau}\right)\right| < \mu \quad \textit{for all } 0 \le \tau \le T. \tag{4.2}$$

Before we give the proof, we make a few remarks:

(i) The only information about a Gaussian beam we made use of in Theorem 2.1, apart from it being an approximate solution, was that it is supported in a given neighbourhood $\mathcal{N}$ of the null geodesic γ. This then yielded, together with (2.2), an estimate on the energy outside of the neighbourhood $\mathcal{N}$ of the actual solution to the wave equation, so we could construct solutions to the wave equation with *localised* energy. However, Theorem 2.1 does not make any statement about the *temporal behaviour* of this localised energy. The above theorem fills this gap by investigating the temporal behaviour of the energy of the approximate solution, i.e., of the Gaussian beam. Together with (2.2) (or even with (2.10)!) this then gives an estimate on the temporal behaviour of the localised energy of the actual solution to the wave equation.

(ii) If N is a timelike Killing vector field, the N-energy $-g(N, \dot{\gamma})$ of the null geodesic γ is constant and, thus, so is approximately the N-energy of the Gaussian beam.

(iii) By our Definition 3.13 a Gaussian beam is a complex-valued function. However, by taking the real or the imaginary part, one can also define a real-valued Gaussian beam. The result of Theorem 4.1 also holds true in this case, and can be proved using exactly the same technique — only the computations become a bit longer, since we have to deal with more terms.

(iv) Although we have stated the above theorem again using the general assumptions needed for Theorem 2.1, we actually do not need more assumptions than we need for the construction of a Gaussian beam; see the final remark of the previous section.

Proof. Recall from Definition 3.13 that a Gaussian beam $\tilde{u}_{\lambda,\mathcal{N}}$ along γ with structure functions a and ϕ, parameters $\mathcal{N}$ and λ, and initial N-energy equal to $-g(N, \dot{\gamma})\big|_{\gamma(0)}$, is a function

$$\tilde{u}_{\lambda,\mathcal{N}} = \frac{u_{\lambda,\mathcal{N}}}{\sqrt{E^N_0(u_{\lambda,\mathcal{N}})}} \cdot \sqrt{-g(N, \dot{\gamma})|_{\gamma(0)}} = \frac{a_{\mathcal{N}} e^{i\lambda\phi}}{\sqrt{E^N_0(u_{\lambda,\mathcal{N}})}} \cdot \sqrt{-g(N, \dot{\gamma})|_{\gamma(0)}},$$

where the functions $a_{\mathcal{N}}$ and ϕ satisfy (3.2), (3.3), (3.9), (3.10), (3.11), (3.12), $\operatorname{supp}(a_{\mathcal{N}}) \subseteq \mathcal{N}$, $\mathcal{N} \cap R_{[0,T]}$ is relatively compact for all $T > 0$ with $\Sigma_T \cap \operatorname{Im}(\gamma) \neq \varnothing$, and, for a cover of γ with slice coordinate patches, (3.4) holds for all $x \in \operatorname{supp}(a_{\mathcal{N}})$.

We will show

$$E_\tau^N(\tilde{u}_{\lambda,\mathcal{N}}) = \frac{E_\tau^N(u_{\lambda,\mathcal{N}})}{E_0^N(u_{\lambda,\mathcal{N}})} \cdot \left[-g(N,\dot{\gamma})\big|_{\gamma(0)}\right] = -g(N,\dot{\gamma})\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau} + o(\lambda), \tag{4.3}$$

where $o(\lambda)$ goes to zero uniformly in $0 \le \tau \le T$ for $\lambda \to \infty$. This would then prove the theorem.

In the following we compute the leading-order term of $E_\tau^N(u_{\lambda,\mathcal{N}})$ in λ:

$$\begin{aligned}
J^N(u_{\lambda,\mathcal{N}})\cdot n_{\Sigma_\tau} &= \mathfrak{Re}(N u_{\lambda,\mathcal{N}} \cdot \overline{n_{\Sigma_\tau} u_{\lambda,\mathcal{N}}}) - \tfrac{1}{2} g(N, n_{\Sigma_\tau})\, du_{\lambda,\mathcal{N}} \cdot \overline{du_{\lambda,\mathcal{N}}} \\
&= \lambda^2 |a_{\mathcal{N}}|^2 N\phi_1 \cdot n_{\Sigma_\tau}\phi_1 \cdot e^{-2\lambda\phi_2} + \lambda^2 |a_{\mathcal{N}}|^2 N\phi_2 \cdot n_{\Sigma_\tau}\phi_2 \cdot e^{-2\lambda\phi_2} + \mathcal{O}(\lambda)\cdot e^{-2\lambda\phi_2} \\
&\qquad - \tfrac{1}{2} g(N, n_{\Sigma_\tau})[\lambda^2 |a_{\mathcal{N}}|^2\,(d\phi_1 \cdot d\phi_1)\, e^{-2\lambda\phi_2} + \lambda^2 |a_{\mathcal{N}}|^2\,(d\phi_2 \cdot d\phi_2)\, e^{-2\lambda\phi_2} + \mathcal{O}(\lambda)\cdot e^{-2\lambda\phi_2}].
\end{aligned}$$

Note that $d\phi_2\big|_{\gamma(\tau)} = 0$, so these terms are of lower order after integration over Σ_τ. The same holds for the $d\phi_1 \cdot d\phi_1$ term. Thus, we get

$$E_\tau^N(u_{\lambda,\mathcal{N}}) = \underbrace{\lambda^2 \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 N\phi_1 \cdot n_{\Sigma_\tau}\phi_1 e^{-2\lambda\phi_2}\, \mathrm{vol}_{\bar{g}_\tau}}_{=\mathcal{O}(\lambda^{1/2})} + \underbrace{\text{lower-order terms}}_{=\mathcal{O}(1)}. \tag{4.4}$$

The main part of the proof is an approximate conservation law. Recall that $a_{\mathcal{N}}$ and ϕ satisfy (3.9) and (3.10). These equations yield

$$\begin{aligned}
\operatorname{grad}\phi(|a_{\mathcal{N}}|^2) &= \operatorname{grad}\phi(a_{\mathcal{N}}) \cdot \overline{a_{\mathcal{N}}} + a_{\mathcal{N}} \cdot \operatorname{grad}\phi(\overline{a_{\mathcal{N}}}) \\
&= -\tfrac{1}{2}(\Box\phi \cdot a_{\mathcal{N}}\overline{a_{\mathcal{N}}} + a_{\mathcal{N}}\,\overline{\Box\phi} \cdot \overline{a_{\mathcal{N}}}) = -\mathfrak{Re}(\Box\phi)|a_{\mathcal{N}}|^2 \quad \text{along } \gamma
\end{aligned} \tag{4.5}$$

and

$$d\phi \cdot d\phi = (d\phi_1 + i d\phi_2)\cdot(d\phi_1 + i d\phi_2) = d\phi_1 \cdot d\phi_1 - d\phi_2 \cdot d\phi_2 + 2i\, d\phi_1 \cdot d\phi_2$$

vanishes to second order along γ; thus, in particular,

$$d\phi_1 \cdot d\phi_2 = \operatorname{grad}\phi_1\,(\phi_2) \quad \text{vanishes along } \gamma \text{ to second order.} \tag{4.6}$$

Lemma 3.6(ii), together with (4.5) and (4.6), shows that the current

$$X_{\lambda,\mathcal{N}} = \lambda^2 \cdot |a_{\mathcal{N}}|^2 e^{-2\lambda\phi_2}\, \operatorname{grad}\phi_1$$

is approximately conserved in the sense that

$$\begin{aligned}
&\int_{R_{[0,\tau]}} \operatorname{div} X_{\lambda,\mathcal{N}}\, \mathrm{vol}_g \\
&\quad = \lambda^2 \cdot \int_{R_{[0,\tau]}} \Big(\underbrace{[\operatorname{grad}\phi_1(|a_{\mathcal{N}}|^2) + \Box\phi_1 \cdot |a_{\mathcal{N}}|^2] e^{-2\lambda\phi_2}}_{=\lambda^{-1/2}\cdot\lambda^{-3/2}=\lambda^{-2} \text{ after integration}} - \underbrace{2\lambda \operatorname{grad}\phi_1(\phi_2)\cdot|a_{\mathcal{N}}|^2 e^{-2\lambda\phi_2}}_{=\lambda\cdot\lambda^{-3/2}\cdot\lambda^{-3/2}=\lambda^{-2} \text{ after integration}}\Big)\, \mathrm{vol}_g = \mathcal{O}(1),
\end{aligned}$$

but

$$\int_{\Sigma_\tau} X_{\lambda,\mathcal{N}} \cdot n_{\Sigma_\tau} \operatorname{vol}_{\bar{g}_\tau} = \lambda^2 \cdot \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_\tau} = \mathcal{O}(\lambda^{1/2}).$$

In particular, we obtain[23]

$$\left|\lambda^2 \cdot \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_\tau} - \lambda^2 \cdot \int_{\Sigma_0} |a_{\mathcal{N}}|^2 n_{\Sigma_0} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_0}\right| = \left|\int_{R_{[0,\tau]}} \operatorname{div} X_{\lambda,\mathcal{N}} \operatorname{vol}_g\right| = \mathcal{O}(1). \quad (4.7)$$

We also observe that, by Lemma 3.6(ii), we have

$$\lambda^2 \cdot \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 \big(N\phi_1 - N\phi_1\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau}\big) \cdot n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_\tau} = \mathcal{O}(1). \quad (4.8)$$

It thus follows from (4.4), (4.7), and (4.8) that

$$\begin{aligned} E^N_\tau(u_{\lambda,\mathcal{N}}) &= \lambda^2 \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 N\phi_1 \cdot n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_\tau} + \mathcal{O}(1) \\ &= \lambda^2 \cdot N\phi_1\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau} \int_{\Sigma_\tau} |a_{\mathcal{N}}|^2 n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_\tau} + \mathcal{O}(1) \\ &= \lambda^2 \cdot N\phi_1\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau} \int_{\Sigma_0} |a_{\mathcal{N}}|^2 n_{\Sigma_\tau} \phi_1 e^{-2\lambda\phi_2} \operatorname{vol}_{\bar{g}_0} + \mathcal{O}(1) \\ &= \frac{N\phi_1\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau}}{N\phi_1\big|_{\operatorname{Im}(\gamma)\cap\Sigma_0}} \cdot E^N_0(u_{\lambda,\mathcal{N}}) + \mathcal{O}(1) \\ &= \frac{g(N,\dot{\gamma})\big|_{\operatorname{Im}(\gamma)\cap\Sigma_\tau}}{g(N,\dot{\gamma})\big|_{\operatorname{Im}(\gamma)\cap\Sigma_0}} \cdot E^N_0(u_{\lambda,\mathcal{N}}) + \mathcal{O}(1). \end{aligned}$$

Substituting this into the expression for $E^N_\tau(\tilde{u}_{\lambda,\mathcal{N}})$, i.e., the first equation in (4.3), we obtain the second equation of (4.3). This finishes the proof of Theorem 4.1. □

5. Some general theorems about the Gaussian beam limit of the wave equation

We can now make a much more detailed statement about the behaviour of solutions v of the wave equation in the Gaussian beam limit than Theorem 2.1 does:

Theorem 5.1. *Let (M, g) be a time-oriented, globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$, where Σ_0 is a Cauchy hypersurface. Furthermore, let $\gamma : [0, S) \to M$ be an affinely parametrised future-directed null geodesic with $\gamma(0) \in \Sigma_0$, where $0 < S \leq \infty$. Finally, let N be a timelike, future-directed vector field.*

For any neighbourhood $\mathcal{N}$ of γ, any $T > 0$ with $\Sigma_T \cap \operatorname{Im}(\gamma) \neq \varnothing$, and any $\mu > 0$, there exists a solution $v \in C^\infty(M, \mathbb{C})$ of the wave equation (1.10) with $E^N_0(v) = -g(N, \dot{\gamma})\big|_{\gamma(0)}$ such that

$$\left|E^N_{\tau,\mathcal{N}\cap\Sigma_\tau}(v) - \big[-g(N,\dot{\gamma})\big|_{\operatorname{Im}\gamma\cap\Sigma_\tau}\big]\right| < \mu \quad \textit{for all } 0 \leq \tau \leq T \quad (5.2)$$

[23] In the geometric optics approximation we have, indeed, a proper conservation law, which is interpreted in the physics literature as conservation of photon number; see, for example, [Misner et al. 1973, Chapter 22.5.]

and[24]

$$E^N_{\tau,\mathcal{N}^c\cap\Sigma_\tau}(v) < \mu \quad \textit{for all } 0 \le \tau \le T \tag{5.3}$$

provided that we have, on $R_{[0,T]} \cap J^+(\mathcal{N} \cap \Sigma_0)$,

$$\frac{1}{|dt(n_{\Sigma_\tau})|} + |g(N, n_{\Sigma_\tau})| \le C < \infty \quad \text{and} \quad 0 < c \le |g(N,N)|,$$
$$|\nabla N(n_{\Sigma_\tau}, n_{\Sigma_\tau})| + \sum_{i=1}^{3} |\nabla N(n_{\Sigma_\tau}, e_i)| + \sum_{i,j=1}^{3} |\nabla N(e_i, e_j)| \le C < \infty, \tag{5.4}$$

where c and C are positive constants and $\{n_{\Sigma_\tau}, e_1, e_2, e_3\}$ is an orthonormal frame.

Moreover, by choosing $\mathcal{N}$ a bit smaller, if necessary, (5.2) *holds independently of* (5.4).

Proof. This follows easily from Theorem 2.1, Theorem 4.1, the second part of Remark 2.9 and the triangle inequality for the square root of the N-energy. □

Let us again remark that the solution v of the wave equation in Theorem 5.1 can also be chosen to be real valued.

The next theorem is a direct consequence of Theorem 5.1 and can be used, in particular, but not only, for proving upper bounds on the rate of the energy decay of waves on globally hyperbolic Lorentzian manifolds if we only allow the initial energy on the right-hand side of the decay statement.

Theorem 5.5. *Let (M, g) be a time-oriented globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$, where Σ_0 is a Cauchy hypersurface. Furthermore, let $\mathcal{T}$ be an open subset of M. Assume there is an affinely parametrised future-directed null geodesic $\gamma : [0, S) \to M$ with $\gamma(0) \in \Sigma_0$, where $0 < S \le \infty$, that is completely contained in $\mathcal{T}$. Let*

$$\tau^* := \sup\big\{\hat\tau \in [0,\infty) \;\big|\; \mathrm{Im}(\gamma) \cap \Sigma_\tau \neq \varnothing \text{ for all } 0 \le \tau < \hat\tau\big\}.$$

Moreover, let N be a timelike, future-directed vector field and $P : [0, \tau^*) \to (0, \infty)$ a function.[25]

If there is no constant $C > 0$ such that

$$-g(N, \dot\gamma)\big|_{\mathrm{Im}(\gamma)\cap\Sigma_\tau} \le P(\tau)C$$

holds for all $0 \le \tau < \tau^*$, then there exists no constant $C > 0$ such that

$$E^N_{\tau,\mathcal{T}\cap\Sigma_\tau}(u) \le P(\tau) C E^N_0(u) \tag{5.6}$$

holds for all solutions u of the wave equation (1.10) *for* $0 \le \tau < \tau^*$.

Proof. Assume the contrary, that is, that there exists a constant $C_0 > 0$ such that (5.6) holds. There is then a $0 \le \tau_0 < \tau^*$ with $-g(N,\dot\gamma)\big|_{\mathrm{Im}(\gamma)\cap\Sigma_{\tau_0}} > -g(N,\dot\gamma)\big|_{\mathrm{Im}(\gamma)\cap\Sigma_0} C_0 P(\tau_0)$. Choosing $\mu > 0$ small enough and a neighbourhood $\mathcal{N} \subseteq \mathcal{T}$ of γ small enough such that (5.2) of Theorem 5.1 applies without reference to (5.4), we obtain a contradiction. □

[24] We denote the complement of $\mathcal{N}$ in M by $\mathcal{N}^c$.

[25] There is no assumption on the regularity of the function P.

A very robust method for proving decay of solutions of the wave equation was given in [Dafermos and Rodnianski 2010b] (but also see [Metcalfe et al. 2012]). This method requires an integrated local energy decay (ILED) statement (possibly with loss of derivative), i.e., a statement of the form (5.8). The next theorem gives a sufficient criterion for an ILED statement having to lose regularity.

Theorem 5.7. *Let (M, g) be a time-oriented, globally hyperbolic Lorentzian manifold with time function t, foliated by the level sets $\Sigma_\tau = \{t = \tau\}$, where Σ_0 is a Cauchy hypersurface. Furthermore, let $\mathcal{T}$ be an open subset of M. Assume there is an affinely parametrised, future-directed null geodesic $\gamma : [0, S) \to M$ with $\gamma(0) \in \Sigma_0$, where $0 < S \leq \infty$, that is completely contained in $\mathcal{T}$. Let N be a timelike, future-directed vector field and set*

$$\tau^* := \sup\bigl\{\hat{\tau} \in [0, \infty) \bigm| \mathrm{Im}(\gamma) \cap \Sigma_\tau \neq \varnothing \text{ for all } 0 \leq \tau < \hat{\tau}\bigr\}.$$

If

$$\int_0^{\tau^*} -g(N, \dot{\gamma})\big|_{\mathrm{Im}(\gamma)\cap\Sigma_\tau} \, d\tau = \infty,$$

where $\dot{\gamma}$ is with respect to some affine parametrisation, then there exists no constant $C > 0$ such that

$$\int_0^{\tau^*} \int_{\Sigma_\tau \cap \mathcal{T}} J^N(u) \cdot n_{\Sigma_\tau} \operatorname{vol}_{\bar{g}_\tau} d\tau \leq C E_0^N(u) \tag{5.8}$$

holds for all solutions u of the wave equation (1.10).

The proof of this theorem goes along the same lines as the one of Theorem 5.5. The reader might have noticed that whether an ILED statement of the form (5.8) exists or not depends heavily on the choice of the time function. On the other hand, it also depends heavily on the choice of the time function whether an ILED statement is helpful or not. So, for instance, we only have an estimate of the form

$$\int_{\mathcal{T}\cap R_{[0,\tau^*]}} J^N(u) \cdot n_{\Sigma_\tau} \operatorname{vol}_g \leq C \cdot \int_0^{\tau^*} \int_{\Sigma_\tau \cap \mathcal{T}} J^N(u) \cdot n_{\Sigma_\tau} \operatorname{vol}_{\bar{g}_\tau} d\tau,$$

where $C > 0$, if the time function t is chosen such that $1/|dt(n_{\Sigma_\tau})| \leq C$ is satisfied for all $0 \leq \tau \leq \tau^*$. Such an estimate, together with an ILED statement, is very convenient whenever one needs to control spacetime integrals that are quadratic in the first derivatives of the field.

Part II. Applications to black hole spacetimes

In the following we give a selection of applications of Theorems 5.1, 5.5 and 5.7. A rich variety of behaviours of the energy is provided by black hole spacetimes arising in general relativity.[26] Although we will briefly introduce the Lorentzian manifolds that represent these black hole spacetimes, the reader completely unfamiliar with those is referred to [Hawking and Ellis 1973] for a more detailed discussion, including the concept of a so called Penrose diagram and an introduction to general relativity.

[26] Another physically interesting application would be, for example, to the study of waves in time-dependent inhomogeneous media.

We first restrict our considerations to the 2-parameter family of Reissner–Nordström black holes, which are exact solutions to the Einstein–Maxwell equations. The spherical symmetry of these spacetimes (and the accompanying simplicity of the metric) allows for an easy presentation without hiding any crucial details. In Section 7 we then discuss the Kerr family and show that analogous results hold.

6. Applications to Schwarzschild and Reissner–Nordström black holes

The 2-parameter family of Reissner–Nordström spacetimes is given by

$$g = -\left(1 - \frac{2m}{r} + \frac{e^2}{r^2}\right) dt^2 + \left(1 - \frac{2m}{r} + \frac{e^2}{r^2}\right)^{-1} dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2, \tag{6.1}$$

initially defined on the manifold $M := \mathbb{R} \times (m + \sqrt{m^2 - e^2}, \infty) \times \mathbb{S}^2$, for which (t, r, θ, φ) are the standard coordinates. We restrict the real parameters m and e, which model the mass and the charge of the black hole, respectively, to the range $0 \le e \le m$, $m \neq 0$.

For $e = 0$ we obtain the 1-parameter Schwarzschild subfamily which solves the vacuum Einstein equations. The manifold M and the metric (6.1) can be analytically extended (so that they still solve the Einstein equations). The so-called Penrose diagram of the maximal analytic extension of the Schwarzschild family is given in Figure 2.

The diamond-shaped region to the right corresponds to the Lorentzian manifold (M, g) we started with; it represents the exterior of the black hole. The triangle to the top corresponds to the interior of the black hole, which is separated from the exterior by the so-called event horizon, the line from the centre to the top-right i^+. The remaining parts of the Penrose diagram play no role in the following discussion.

The black hole stability problem (see the introduction of [Dafermos and Rodnianski 2013]) motivates the study of the wave equation in the exterior of the black hole (the event horizon included). In accordance with our discussion in Section 1B, we consider the framework of the energy method for the study of the wave equation. A suitable notion of energy for the black hole exterior is obtained via (1.9) through the

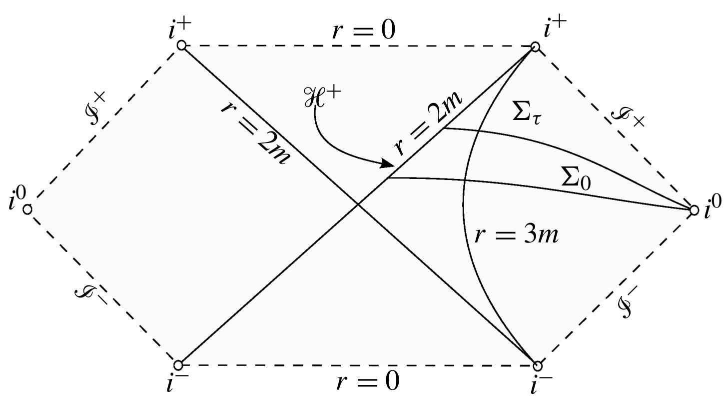

Figure 2. The Penrose diagram of the maximal analytic extension of the Schwarzschild family.

foliation given by $\Sigma_\tau = \{t^* = \tau\}$ for $t^* \geq c > -\infty$, where $t^* = t + 2m\log(r-2m)$, together with the timelike vector field $N := -(dt^*)^\sharp$.[27]

6A. *Trapping at the photon sphere.* There are null geodesics in the Schwarzschild spacetime that stay forever on the photon sphere at $\{r = 3m\}$. Indeed, one can check that the curve γ given by

$$\gamma(s) = \big(s, 3m, \tfrac{1}{2}\pi, (27m^2)^{-1/2}s\big)$$

in (t, r, θ, φ) coordinates is an affinely parametrised null geodesic with N-energy given by $-g(N, \dot\gamma) = 1$. We now apply Theorem 5.5: The time-oriented,[28] globally hyperbolic Lorentzian manifold can be taken to be the domain of dependence $\mathscr{D}(\Sigma_0)$ of Σ_0 in (M, g). Moreover, we choose the time function to be given by the restriction of t^* to $\mathscr{D}(\Sigma_0)$, and the vector field N and null geodesic $\gamma(s)$ in Theorem 5.5 are given by N and $\gamma(s - 2m\log m)$ from above. Since $-g(N, \dot\gamma) = 1$ holds, Theorem 5.5 now states that, given any open neighbourhood $\mathcal{T}$ of $\mathrm{Im}(\gamma)$ in $\mathscr{D}(\Sigma_0)$, there is no function $P : [0, \infty) \to (0, \infty)$ with $P(\tau) \to 0$ for $\tau \to \infty$ such that

$$E^N_{\tau, \mathcal{T}\cap\Sigma_\tau}(u) \leq P(\tau) E^N_0(u)$$

holds for all solutions u of the wave equation for all $\tau \geq 0$. It follows that an LED statement for such a region can only hold if it loses differentiability. One can infer the analogous result about ILED statements from Theorem 5.7.

Let us mention here that γ has conjugate points. Indeed, the Jacobi field J with initial data $J(0) = 0$ and $D_s J(0) = \partial_\theta|_{\gamma(0)}$ vanishes in finite affine time $s > 0$: First note that the vector field

$$s \mapsto \partial_\theta\big|_{\gamma(s)}$$

along γ is parallel, i.e., $D_s \partial_\theta\big|_{\gamma(s)} = 0$. Moreover, a direct computation yields

$$R(\partial_\theta, \dot\gamma)\dot\gamma\big|_{\gamma(s)} = \frac{1}{27m^2}\partial_\theta\Big|_{\gamma(s)},$$

where $R(\cdot\,,\cdot)\,\cdot$ is the Riemann curvature endomorphism. Thus, it follows that the vector field

$$J(s) = (27m^2)^{1/2}\sin((27m^2)^{-1/2}s)\cdot\partial_\theta\big|_{\gamma(s)}$$

[27] We are intentionally quite vague about what we mean by "suitable notion of energy". Instead of considering a foliation that ends at spacelike infinity ι^0, it is sometimes desirable to work with a foliation that ends at future null infinity $\mathcal{I}^+$. In a stationary spacetime, however, it is always convenient (and indeed "suitable") to work with a foliation and an energy-measuring vector field N both of which are invariant under the flow of the Killing vector field. The obvious advantage is that the constants in Sobolev embeddings do not depend on the leaf — provided, of course, that higher-energy norms are also defined accordingly. The precise choice of the timelike vector field N in a *compact* region of one leaf is completely irrelevant, since all the energy norms are equivalent in a compact region. In particular, one can deduce that the following result about trapping at the photon sphere in Schwarzschild remains unchanged if we choose a different timelike vector field N which commutes with ∂_t and a different foliation by spacelike slices. In fact, note that the behaviour of the energy of the null geodesic, $-g(N, \dot\gamma)$, does not depend at all on the choice of the foliation!

[28] The time orientation is given by the timelike vector field N.

satisfies the Jacobi equation $D_t^2 J + R(J, \dot{\gamma})\dot{\gamma} = 0$. Moreover, it clearly satisfies the above initial conditions and vanishes in finite affine time.

It now follows from Theorem A.1 that one cannot construct localised solutions to the wave equation along the trapped null geodesic γ using the naive geometric optics approximation alone. Indeed, one would need to bridge these caustics using Maslov's canonical operator.

That one can indeed prove an (I)LED statement with a loss of derivative was shown in [Dafermos and Rodnianski 2009] (see also [Blue and Sterbenz 2006]). In fact, it is sufficient to lose only an ε of a derivative; see [Blue and Soffer 2009] and also [Dafermos and Rodnianski 2013]. For a numerical study of the behaviour of a wave trapped at the photon sphere we refer the interested reader to [Zenginoglu and Galley 2012].

Other, similar, examples are trapping at the photon sphere in higher-dimensional Schwarzschild [Schlue 2013] or in Reissner–Nordström [Aretakis 2011a; Blue and Soffer 2009].

6B. ***The red-shift effect at the event horizon — and its relevance for scattering constructions from the future.*** Another kind of behaviour of the energy is exhibited by the trapping occuring at the event horizon of the Schwarzschild spacetime. Recall that the event horizon $\mathcal{H}^+$ at $\{r = 2m\}$ is a null hypersurface, spanned by null geodesics. In $(t^*, r, \theta, \varphi)$ coordinates the affinely parametrised generators are given by

$$\gamma(s) = \Big(\frac{1}{\kappa}\log s, 2m, \theta_0, \varphi_0\Big),$$

where $\kappa = 1/(4m)$ is the surface gravity, $s \in (0, \infty)$ and θ_0, φ_0 are constants. Thus, we have

$$-(\dot{\gamma}(s), N) = \frac{1}{\kappa s} = \frac{1}{\kappa} e^{-\kappa t^*}, \tag{6.2}$$

i.e., the energy of the corresponding Gaussian beam decays exponentially — a direct manifestation of the celebrated red-shift effect. For more on the impact of the red-shift effect on the study of the wave equation on Schwarzschild we refer the reader to the original paper by Dafermos and Rodnianski [2009], but also see [Dafermos and Rodnianski 2013].

Let us emphasise again that the null geodesics at the photon sphere as well as those at the horizon are trapped, in the sense that they never escape to null infinity — but only those at the photon sphere form an obstruction for an LED statement without loss of differentiability; the "trapped" energy at the horizon decays exponentially. This is in stark contrast to the obstacle problem, where every trapped light ray automatically leads to an obstruction for an LED statement without loss of derivatives (see [Ralston 1969]). This new variety of how the "trapped" energy behaves is due to the lack of a global timelike Killing vector field.

Let us now investigate the role played by the red-shift effect in scattering constructions from the future. While the red-shift effect is conducive to proving bounds on solutions to the wave equation in the "forward problem", it turns into a blue-shift in the "backwards problem" (see Figure 3);[29] it amplifies energy near the horizon.

[29] We call the initial value problem on Σ_0 to the future the "forward problem", while solving a mixed characteristic initial value problem on $\mathcal{H}^+(\tau) \cup \Sigma_\tau$ to the past (or indeed a scattering construction from the future with data on $\mathcal{H}^+$ and $\mathcal{I}^+$) is called the "backwards problem". Here, we have denoted the (closed) portion of the event horizon $\mathcal{H}^+$ that is cut out by Σ_0 and Σ_τ by $\mathcal{H}^+(\tau)$.

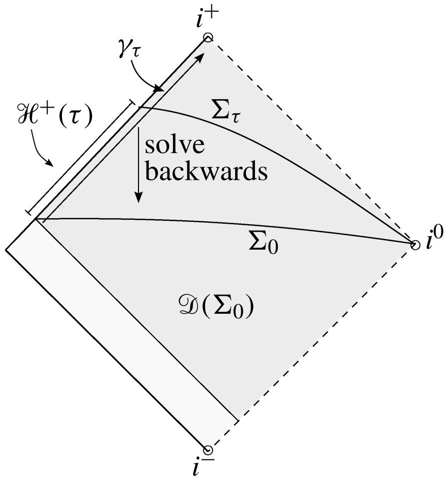

Figure 3. Illustration of the backwards problem.

Proposition 6.3. *For every $\mu > 0$ and every $\tau > 0$ there exists a smooth solution*[30] $v \in C^\infty(\overline{\mathcal{D}(\Sigma_0)}, \mathbb{C})$ *to the wave equation (1.10) with $E^N_\tau(v) = 1$ and $\int_{\mathcal{H}^+(\tau)} J^N(v) \lrcorner \operatorname{vol}_g < \mu$, which satisfies $E^N_0(v) \geq e^{\kappa\tau} - \mu$, where $\kappa = 1/(4m)$ is the surface gravity of the Schwarzschild black hole.*

Here, $J^N(v) \lrcorner \operatorname{vol}_g$ denotes the 3-form obtained by inserting the vector field $J^N(v)$ into the first slot of vol_g. Let us also remark that μ should be thought of as a small positive number, while τ rather as a big one.

Proof. As in Section 6A, we consider the Lorentzian manifold $\mathcal{D}(\Sigma_0)$ with time function t^* and timelike vector field N. Since geodesics depend smoothly on their initial data, it follows from (6.2) that we can find, for every $\tau > 0$, an affinely parametrised, radially outgoing null geodesic[31] γ_τ in $\mathcal{D}(\Sigma_0)$ with $\big|-(\dot\gamma_\tau, N)|_{\operatorname{Im}(\gamma_\tau)\cap\Sigma_0} - e^{\kappa\tau}\big| < \mu/2$ and $-(\dot\gamma_\tau, N)|_{\operatorname{Im}(\gamma_\tau)\cap\Sigma_\tau} = 1$. We note that, for our choice of time function and vector field N, the condition (2.3) is satisfied, which does not only give us the energy estimate (2.8) but here also the refined version

$$\int_{\mathcal{H}^+(\tau)} J^N(u) \lrcorner \operatorname{vol}_g + E^N_\tau(u) \leq C(\tau)\big(E^N_0(u) + \|\Box u\|^2_{L^2(R_{[0,T]})}\big), \tag{6.4}$$

which holds in $\overline{\mathcal{D}(\Sigma_0)}$ for all $\tau > 0$ and all $u \in C^\infty(\overline{\mathcal{D}(\Sigma_0)}, \mathbb{R})$. The estimate (6.4) is derived in the same way as (2.8), namely by an application of Stokes' theorem to $J^N(u) \lrcorner \operatorname{vol}_g$, followed by Gronwall's inequality. The estimate (6.4) gives, in addition to (2.2) in Theorem 2.1, the estimate

$$\int_{\mathcal{H}^+(\tau)} J^N(v - \tilde u) \lrcorner \operatorname{vol}_g < \mu, \tag{6.5}$$

where $\tilde u$ is the Gaussian beam and v is the actual solution, as in Theorem 2.1. We now apply Theorem 5.1, where the Lorentzian manifold is given by $\mathcal{D}(\Sigma_0)$, the time function by t^*, the timelike vector field

[30] We denote by $\overline{\mathcal{D}(\Sigma_0)}$ the closure of $\mathcal{D}(\Sigma_0)$ in the maximal analytic extension of Schwarzschild; see Figure 2 on page 1399.

[31] Radially outgoing null geodesics are the lines parallel to, and to the right of, $\mathcal{H}^+$ in the Penrose diagram. In (u, r, θ, φ) coordinates, where $u(t, r, \theta, \varphi) := t - 2m\log(r - 2m) - r$, these null geodesics are tangent to $\partial/\partial r$.

by N and, for given $\tau > 0$, the affinely parametrised null geodesic is taken to be γ_τ from above. For our purposes we can choose any neighbourhood $\mathcal{N}$ of $\mathrm{Im}(\gamma_\tau)$ in $\mathcal{D}(\Sigma_0)$. Theorem 5.1 then ensures the existence of a solution $v \in C^\infty(\mathcal{D}(\Sigma_0), \mathbb{C})$ to the wave equation with $E_0^N(v) \geq e^{\kappa\tau} - \mu$ and $E_\tau^N(v) = 1$ — possibly after renormalising the energy at time τ of v to be exactly 1. It is not difficult to show, for example by considering the Cauchy problem for a slightly larger globally hyperbolic Lorentzian manifold which contains the event horizon, that v can be chosen to extend smoothly to the event horizon. We then obtain $\int_{\mathcal{H}^+(\tau)} J^N(v) \lrcorner \mathrm{vol}_g < \mu$ from (6.5), since we recall that the Gaussian beam $\tilde{u}$ in Theorem 2.1 is supported in $\mathcal{N}$, which is disjoint from $\mathcal{H}^+$. This finishes the proof. □

The above proposition shows that for every $\tau > 0$ one can prescribe initial data for the mixed characteristic initial value problem on $\mathcal{H}^+ \cup \Sigma_\tau$ so that the total initial energy is equal to one, while the energy of the solution obtained by solving backwards grows exponentially to $\approx e^{\kappa\tau}$ on Σ_0. Dafermos, Holzegel and Rodnianski [Dafermos et al. 2013] approach the scattering problem from the future for the Einstein equations (with initial data prescribed on $\mathcal{H}^+$ and $\mathcal{I}^+$) by considering it as the limit of finite backwards problems, which — for the wave equation — are qualitatively the same as the backwards problem with initial data on $H^+(\tau)$ and Σ_τ. In order to take the limit of the finite problems, uniform control over the solutions is required: Dafermos et al. use a backwards energy estimate which bounds the energy on Σ_0 by the initial energy on $\mathcal{H}^+$ and Σ_τ, multiplied by $C \cdot \exp(c\tau)$, where c and C are constants that are independent of τ. Proposition 6.3 shows now that this estimate is sharp, in the sense that one cannot avoid exponential growth (at least not as long as one does not sacrifice regularity in the estimate). In particular, working with this estimate enforces the assumption of exponential decay on the scattering data in [Dafermos et al. 2013].

6C. *The blue-shift near the Cauchy horizon of a subextremal Reissner–Nordström black hole.* We now move on to the subextremal Reissner–Nordström black hole, i.e., to the parameter range $0 < e < m$ in (6.1). More precisely, we consider again its maximal analytic extension. Part of the Penrose diagram is given in Figure 4.

Again, the diamond-shaped region *I* represents the black hole exterior and corresponds to the Lorentzian manifold on which the metric g from (6.1) was initially defined. The regions *II*, *III* and *IV* represent the black hole interior. Recall that Reissner–Nordström is a spherically symmetric spacetime. The "radius" of the spheres of symmetry is given by a globally defined function r. We write $D(r) := 1 - 2m/r + e^2/r^2$ and denote the two roots of D by $r_\pm = m \pm \sqrt{m^2 - e^2}$. The future Cauchy horizon[32] is given by $r = r_-$. The coordinate functions (θ, φ) parametrise the spheres of symmetry in the usual way and are globally defined up to one meridian. Regions *I–III* are covered by a (v, r, θ, φ) coordinate chart; in the region I, the function v is given by $v = t + r_I^*$, where r_I^* is a function of r satisfying $dr_I^*/dr = 1/D$. With respect to these coordinates, the Lorentzian metric takes the form

$$g = -D\, dv^2 + dv \otimes dr + dr \otimes dv + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2.$$

[32] We consider a Cauchy surface Σ_0 of the big diamond-shaped region as shown in Figure 5, i.e., a Cauchy surface of the region pictured in Figure 4 without the regions *III* and *IV*.

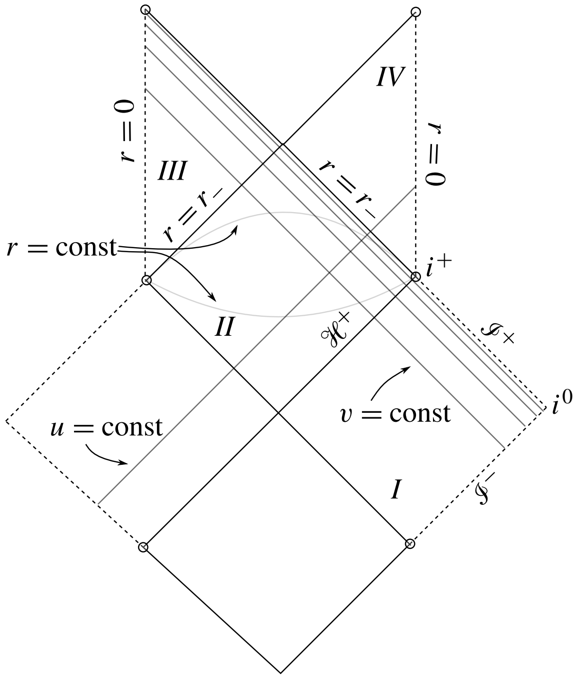

Figure 4. Part of the Penrose diagram for the subextremal Reissner–Nordström black hole.

Introducing a function r^*_{II} in region II, which satisfies $dr^*_{II}/dr = 1/D$ in this region, and defining a function[33] $t := v - r^*_{II}$, we obtain a (t, r, θ, φ) coordinate system for region II in which the metric g is again given by the algebraic expression (6.1). The regions II and IV are covered by a coordinate system (u, r, θ, φ), where the function u is given in region II by $u = t - r^*_{II}$.

Having laid down the coordinate functions we work with, we now investigate the family of affinely parametrised ingoing null geodesics, given in (v, r, θ, φ) coordinates by

$$\gamma_{v_0}(s) = (v_0, -s, \theta_0, \varphi_0),$$

where $s \in (-\infty, 0)$ and we keep θ_0, φ_0 fixed. Clearly, we have[34] $\dot\gamma_{v_0} = -\partial/\partial r\big|_v$. We are interested in the energy of these null geodesics in region II close to i^+ (in the topology of the Penrose diagram), i.e., close to the Cauchy horizon separating region II from region IV. A suitable notion of energy is given by a regular vector field that is future-directed timelike in a neighbourhood of i^+. In order to construct such a vector field, we consider (u, v, θ, φ) coordinates in region II. A straightforward computation shows that

$$N := -\frac{1}{r_+ - r}\frac{\partial}{\partial u}\bigg|_v + \frac{1}{r - r_-}\frac{\partial}{\partial v}\bigg|_u = -\frac{1}{r_+ - r}\frac{\partial}{\partial u}\bigg|_r - \frac{1}{2r^2}(r_+ - r_-)\frac{\partial}{\partial r}\bigg|_u = \frac{r_- - r_+}{2r^2}\frac{\partial}{\partial r}\bigg|_v + \frac{1}{r - r_-}\frac{\partial}{\partial v}\bigg|_r$$

[33]One could also assign the functions t an index, specifying in which region they are defined. Note that these different functions t do *not* patch together to give a globally defined smooth function!

[34]Let us denote with a subscript on the partial derivative which other coordinate (apart from θ and φ) remains fixed.

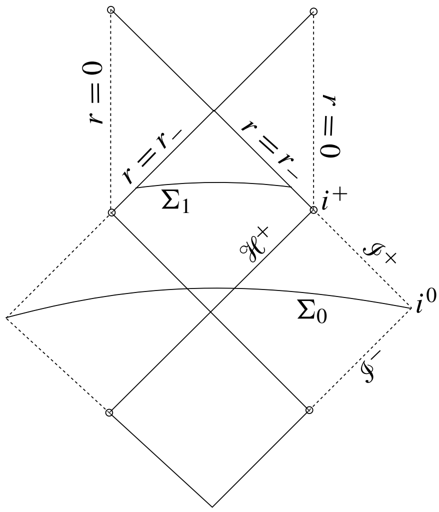

Figure 5. The spacelike slices Σ_0 and Σ_1 of Figure 4.

is future-directed timelike in a neighbourhood of i^+ intersected with region *II* and can be extended to a smooth timelike vector field defined on a neighbourhood of i^+. We obtain

$$-(N, \dot{\gamma}_{v_0}) = \frac{1}{r-r_-}; \tag{6.6}$$

the N-energy of the null geodesics γ_{v_0} gets infinitely blue-shifted near the Cauchy horizon.

For later reference let us note that the rate with which the N-energy (6.6) of γ_{v_0} blows up along a hypersurface of constant u, in advanced time v, is exponential. This is seen as follows: One has

$$r^*_{II}(r) = r + \frac{1}{2\kappa_+}\log(r_+ - r) + \frac{1}{2\kappa_-}\log(r-r_-) + \text{const},$$

where $\kappa_\pm = (r_\pm - r_\mp)/(2r_\pm^2)$ are the surface gravities of the event and the Cauchy horizon, respectively. Thus, for large r^*_{II}, one has $(1/(r-r_-))(r^*_{II}) \sim e^{-2\kappa_- r^*_{II}}$. Finally, along $\{u = u_0 = \text{const}\}$, we have $r^*_{II}(v) = \frac{1}{2}(v-u_0)$. It thus follows that the N-energy (6.6) of γ_{v_0} blows up like $e^{-\kappa_- v}$ along a hypersurface of constant u.

Let us now consider spacelike slices Σ_0 and Σ_1 as in Figure 5, where Σ_0 asymptotes to a hypersurface of constant t and Σ_1 is extendible as a smooth spacelike slice into the neighbouring regions.

Since the normal n_{Σ_1} of Σ_1 is also regular at the Cauchy horizon, it follows from (6.6) that the n_{Σ_1}-energy of the null geodesics γ_{v_0} blows up along Σ_1 when approaching the Cauchy horizon. Moreover, note that the n_{Σ_0}-energy of the geodesics γ_{v_0} along Σ_0 is uniformly bounded as $v_0 \to \infty$. We now apply Theorem 5.1 to the family of null geodesics γ_{v_0} with the following further input: the Lorentzian manifold is given by the domain of dependence $\mathscr{D}(\Sigma_0)$ of Σ_0, the time function is such that Σ_0 and Σ_1 are level sets, N is a timelike vector field that extends n_{Σ_0} and n_{Σ_1}, and finally $\mathcal{N}$ is a small enough neighbourhood of γ_{v_0}. This yields:

Theorem 6.7. *Let Σ_0 and Σ_1 be spacelike slices in the subextremal Reissner–Nordström spacetime as indicated in Figure 6. Then there exists a sequence $\{u_i\}_{i\in\mathbb{N}}$ of solutions to the wave equation with initial energy $E_0^{n_{\Sigma_0}}(u_i) = 1$ on Σ_0 such that the n_{Σ_1}-energy on Σ_1 goes to infinity, i.e., $E_1^{n_{\Sigma_1}}(u_i) \to \infty$ as $i \to \infty$.*

We can infer from Theorem 6.7 that there is no uniform energy boundedness statement — that is, there is no constant $C > 0$ such that

$$\int_{\Sigma_1} J^{n_{\Sigma_1}}(u) \cdot n_{\Sigma_1} \leq C \int_{\Sigma_0} J^{n_{\Sigma_0}}(u) \cdot n_{\Sigma_0} \tag{6.8}$$

holds for all solutions u of the wave equation.

Let us remark here that the nonexistence of a uniform energy boundedness statement has, in particular, the following consequence: one cannot choose a time function for the region bounded by Σ_0 and Σ_1 for which these hypersurfaces are level sets and, moreover, extend the normals of Σ_0 and Σ_1 to a smooth timelike vector field N *in such a way that* an energy estimate of the form (2.8) holds. This emphasises the importance of the condition (2.3) for the *global* approximation scheme on general Lorentzian manifolds and points out the necessity of a *local* understanding of the approximate solution provided by Theorems 4.1 and 5.1.

One actually expects that there is no energy boundedness statement at all, no matter how many derivatives one loses or whether one restricts the support of the initial data:

Conjecture 6.9. *For generic compactly supported smooth initial data on Σ_0, the n_{Σ_1}-energy along Σ_1 of the corresponding solution to the wave equation is infinite.*

Let us remark here that the analysis carried out in [Dafermos 2005] shows in particular that proving the above conjecture can be reduced to proving a lower bound on the decay rate of the spherical mean of the generic solution (as in Conjecture 6.9) on the horizon.

Before we elaborate in Section 6E on the mechanism that leads to the blow-up of the energy near the Cauchy horizon in Theorem 6.7, let us investigate the situation for *extremal* Reissner–Nordström black holes.

6D. ***The blue-shift near the Cauchy horizon of an extremal Reissner–Nordström black hole.*** The extremal Reissner–Nordström black hole is given by the choice $m = e$ of the parameters in (6.1). We again consider the maximal analytic extension of the initially defined spacetime. Part of the Penrose diagram is given in Figure 6.

The region I represents again the black hole exterior and corresponds to the Lorentzian manifold on which the metric g from (6.1) was initially defined. The black hole interior extends over the regions II and III. The discussion of the functions r, θ and φ carries over from the subextremal case. However, in the extremal case, $D(r)$ has a double zero at $r = m$, the value of the radius of the spheres of symmetry on the event, as well as on the Cauchy horizon. The regions I and II can be covered by "ingoing" null coordinates (v, r, θ, φ), where the function v is given in region I by $v = t + r_I^*$, where again $r_I^*(r)$ satisfies $dr_I^*/dr = 1/D$. In the same way as in the subextremal case, one introduces r_{II}^* and defines a (t, r, θ, φ) coordinate system for the region II. Finally, the regions II and III are covered by "outgoing" null coordinates (u, r, θ, φ), where we have $u = t - r_{II}^*$ in region II.

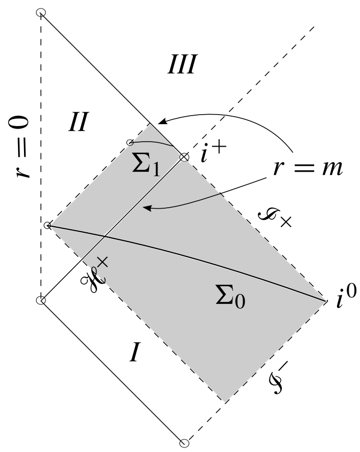

Figure 6. Part of the Penrose diagram for the extremal Reissner–Nordström black hole.

In ingoing null coordinates, the affinely parametrised, radially ingoing null geodesics are given by $\gamma_{v_0}(s) = (v_0, -s, \theta_0, \varphi_0)$, where $s \in (-\infty, 0)$. Expressing the tangent vector of γ_{v_0} in region *II* in outgoing coordinates, we obtain

$$\dot{\gamma}_{v_0} = -\frac{\partial}{\partial r}\Big|_v = \frac{2}{D}\frac{\partial}{\partial u}\Big|_r - \frac{\partial}{\partial r}\Big|_u, \tag{6.10}$$

which blows up at $r = m$. Thus, we have, for any future-directed timelike vector field N in region *II* which extends to a regular timelike vector field in region *III*, that the N-energy $-g(\dot{\gamma}_{v_0}, N)$ of γ_{v_0} blows up along the hypersurface Σ_1 for $v_0 \to \infty$. Choosing now a spacelike slice Σ_0 as in the above diagram, again asymptoting to a $\{t = \text{const}\}$ hypersurface at i^0, and restricting consideration to its domain of dependence, we obtain a globally hyperbolic spacetime (the shaded region) with respect to which we can apply Theorem 5.1, inferring the analogue of Theorem 6.7 for extremal Reissner–Nordström black holes.

For the discussion in the next section, we again investigate the rate, in advanced time v, with which the N-energy $-g(\dot{\gamma}_{v_0}, N)$ blows up along a hypersurface of constant u; here, we have

$$r^*_{II}(r) = r + m\log((r-m)^2) - \frac{m^2}{(r-m)} + \text{const}.$$

It follows that for large r^*_{II} one has $(1/D)(r^*_{II}) \sim (r^*_{II})^2$. Moreover, along $\{u = u_0 = \text{const}\}$, we have $r^*_{II}(v) = \frac{1}{2}(v - u_0)$, from which it follows that the N-energy $-g(\dot{\gamma}_{v_0}, N)$ of the family of null geodesics γ_{v_0} blows up like v^2.

6E. ***The strong and the weak blue-shift — and their relevance for strong cosmic censorship.*** In the example of subextremal Reissner–Nordström as well as in the example of extremal Reissner–Nordström, the energy of the Gaussian beams is blue-shifted near the Cauchy horizon. Although not important for the proof of the *qualitative* result of Theorem 6.7 (and the analogous statement for the extremal case), the difference in the *quantitative* blow-up rate of the energy in the two cases is conspicuous.

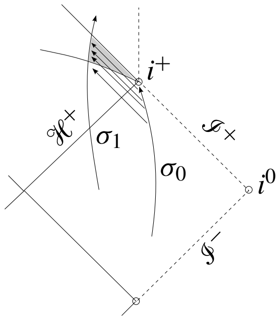

Figure 7. Illustration of the blue-shift effect in subextremal Reissner–Nordström.

Let us first recall the familiar heuristic picture that explains the basic mechanism responsible for the blue-shift effect in both cases;[35] see Figure 7. The observer σ_0 travels along a timelike curve of infinite proper time to i^+ and, in regular time intervals, sends signals of the same energy into the black hole. These signals are received by the observer σ_1, who travels into the black hole and crosses the Cauchy horizon, within *finite* proper time — which leads to an infinite blue shift. This mechanism was first pointed out by Roger Penrose [1968, page 222].[36] Although the picture, along with its heuristics, allow for inferring the *presence* of a blue-shift near the Cauchy horizon, they do not reveal the *strength* of the blue-shift. For investigating the latter, it is important to note that the region in spacetime which actually causes the blue shift is a neighbourhood of the Cauchy horizon. This neighbourhood is not well defined, however, one could think of it as being given by a neighbourhood of constant r — the shaded region in the diagram of subextremal Reissner–Nordström in Figure 7. The crucial difference between the subextremal and the extremal case is that, in the extremal case, the blue-shift degenerates *at the Cauchy horizon itself*, while, in the subextremal case, it does not: the subextremal Cauchy horizon continues to blue-shift radiation. In particular, one can prove an analogous result to Proposition 6.3 there — but for the forward problem.

This degeneration of the blue-shift towards the Cauchy horizon in the extremal case leads to the (total) blue-shift being weaker than the blue-shift in the subextremal case. Thus, the geometry of spacetime near the Cauchy horizon is crucial for understanding the strength of the blue-shift effect, and hence the blow-up rate of the energy.

We now continue with a heuristic discussion of the importance of the different blow-up rates. The reader might have noticed that we only made Conjecture 6.9 for the subextremal case; and indeed, the analogous conjecture for the extremal case is expected to be false: While in our construction we consider

[35] In Figure 7 we give the picture for the subextremal case. However, the picture and the heuristics for the extremal case are exactly the same!

[36] There, he describes the above scenario in the following, more dramatic language (he considers the scenario of gravitational collapse, where the Einstein equations are coupled to some matter model and denotes the Cauchy horizon with $H_+(\mathscr{H})$):

> There is a further difficulty confronting our observer who tries to cross $H_+(\mathscr{H})$. As he looks out at the universe that he is "leaving behind", he sees, in one final flash, as he crosses $H_+(\mathscr{H})$, the entire later history of the rest of his "old universe". [...] If, for example, an unlimited amount of matter eventually falls into the star then presumably he will be confronted with an infinite density of matter along "$H_+(\mathscr{H})$". Even if only a finite amount of matter falls in, it may not be possible, in generic situations to avoid a curvature singularity in place of $H_+(\mathscr{H})$.

a family of ingoing wave packets whose energy along a fixed outgoing null ray to $\mathcal{I}^+$ does *not* decay in advanced time v, the scattered "ingoing energy" of a wave with initial data as in Conjecture 6.9 will decay in advanced time v along such an outgoing null ray. Thus, the blow-up of the energy near the Cauchy horizon can be counteracted by the decay of the energy of the wave towards null infinity. In the extremal case, the blow-up rate is v^2, which does not dominate the decay rate of the energy towards null infinity; the exponential blow-up rate $e^{-\kappa_- v}$, however, does. These are the heuristic reasons for only formulating Conjecture 6.9 for the subextremal case. We conclude with a couple of remarks: Firstly, one should actually compare the decay rate of the ingoing energy not along an outgoing null ray to $\mathcal{I}^+$, but along the event horizon — or even better, along a spacelike slice in the interior of the black hole approaching i^+ in the topology of the Penrose diagram. Secondly, we would like to repeat and stress the point made, namely that the heuristics given in the very beginning of this section, which solely ensure the *presence* of a blue-shift, are not sufficient to cause a C^1 instability of the wave at the Cauchy horizon. For this to happen, the local geometry of the Cauchy horizon is crucial. Finally, let us conjecture, based on the fact that in the extremal case one gains powers of v in the blow-up rate at the Cauchy horizon when considering higher-order energies, that there is some natural number $k > 1$ such that waves with initial data as in Conjecture 6.9 exhibit a C^k instability at the Cauchy horizon.

We conclude this section by recalling that the study of the wave equation on black hole backgrounds serves as a source of intuition for the behaviour of gravitational perturbations of these spacetimes. Thus, the following expected picture emerges: Consider a *generic* dynamical spacetime which at late times approaches a subextremal Reissner–Nordström black hole. Then the Cauchy horizon is replaced by a weak null curvature singularity (for this notion see [Dafermos 2005]).

If we restrict consideration to the class of dynamical spacetimes which at late times approach an extremal Reissner–Nordström black hole, then the *generic* spacetime *within this class* has a more regular Cauchy horizon, which in particular is not seen as a singularity from the point of view of the low regularity well-posedness theory for the Einstein equations; see the resolution [Klainerman et al. 2013] of the L^2-curvature conjecture. This picture is also supported by the recent numerical work [Murata et al. 2013].

6F. ***Trapping at the horizon of an extremal Reissner–Nordström black hole.*** We again consider the extremal Reissner–Nordström black hole. With v defined as in Section 6D, we introduce the function $t^* := v - r$. In the coordinates $(t^*, r, \theta, \varphi)$ the metric takes the form

$$g = -D(dt^*)^2 + (1-D)(dt^* \otimes dr + dr \otimes dt^*) + (2-D)\, dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2.$$

We see that the foliation of the exterior given by $\Sigma_\tau = \{t^* = \tau\}$ is a foliation by spacelike slices, which is invariant under the flow of the stationary Killing vector field ∂_{t^*} and is regular at the event horizon $\mathcal{H}^+$ in the sense that it extends smoothly as a spacelike foliation across the event horizon; see Figure 8.

An appropriate choice of timelike vector field for measuring the energy of waves in the black hole exterior is thus given by $N = -(dt^*)^\sharp$, since it is also invariant under the flow of the Killing vector field ∂_{t^*} and extends smoothly as a timelike vector field across the event horizon. Hence, the corresponding N-energy is nondegenerate at the event horizon. These choices of foliation and timelike vector field N correspond qualitatively to the choices made in the Schwarzschild spacetime in Sections 6A and 6B.

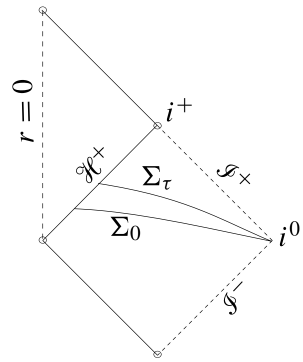

Figure 8. The extremal Reissner–Nordström black hole.

Aretakis [2011a; 2011b] investigated the behaviour of waves on this spacetime and obtained stability (i.e., boundedness and decay results) as well as instability results (blow-up of certain higher-order derivatives along the horizon); for further developments see also [Lucietti and Reall 2012]. The instability results originate from a conservation law on the extremal horizon once decay results for the wave are established. In order to obtain these stability results, Aretakis followed the new method introduced by Dafermos and Rodnianski [2010b].[37] The first important step is to prove an ILED statement. As in the Schwarzschild spacetime we have trapping at the photon sphere (here at $\{r = 2m\}$), and as shown before, an ILED statement has to degenerate there in order to hold. The fundamentally new difficulty in the extremal setting arises from the degeneration of the red-shift effect at the horizon $\mathscr{H}^+$, which was needed for proving an ILED statement that holds up to the horizon (see for example [Dafermos and Rodnianski 2013]). And indeed, the energy of the generators of the horizon is no longer decaying: In $(t^*, r, \theta, \varphi)$ coordinates, the affinely parametrised generators are given by

$$\gamma(s) = (s, m, \theta_0, \varphi_0),$$

where $s \in (-\infty, \infty)$ and again θ_0, φ_0 are fixed. Hence, we see that the N-energy of the generators of the horizon is constant: $-(N, \dot\gamma) = 1$.

If we consider a globally hyperbolic subset of the depicted part of extremal Reissner–Nordström that contains the horizon $\mathscr{H}^+$, for example by extending Σ_0 a bit through the event horizon and then considering its domain of dependence, we can directly infer from Theorems 5.5 and 5.7, by applying them to the null geodesic γ from above, that every (I)LED statement which concerns a neighbourhood of the horizon necessarily has to lose differentiability. However, we can also infer the same result for the wave equation on the Lorentzian manifold $\mathscr{D}(\Sigma_0)$, where “a neighbourhood of the horizon” is “a neighbourhood of the horizon in the previous, bigger spacetime, intersected with $\mathscr{D}(\Sigma_0)$”: Analogous to the proof of Proposition 6.3, we consider a sequence of radially outgoing null geodesics in $\mathscr{D}(\Sigma_0)$ whose initial data on Σ_0 converges to the data of γ from above. For every “neighbourhood of the horizon”, every $\tau_0 > 0$ and every (small) $\mu > 0$, there is then an element γ_0 of the sequence such that

[37] Though in addition he had to work with a degenerate energy, which makes things more complicated.

$-(N, \dot{\gamma}_0)|_{\mathrm{Im}(\gamma_0)\cap\Sigma_\tau} \in (1-\mu, 1+\mu)$ for all $0 \le \tau \le \tau_0$. This follows again from the smooth dependence of geodesics on their initial data. We now apply Theorem 5.1 to this sequence of null geodesics to infer that, for every "neighbourhood of the horizon" and for every $\tau_0 > 0$, we can construct a solution to the wave equation whose energy in this neighbourhood is, say, bigger than $\frac{1}{2}$ for all times τ with $0 \le \tau \le \tau_0$. This proves again that there is no nondegenerate (I)LED statement concerning "a neighbourhood of the horizon" in $\mathscr{D}(\Sigma_0)$; the trapping at the event horizon obstructs local energy decay — which is in stark contrast to subextremal black holes.

One should ask now whether an ILED statement with loss of derivative can actually hold. To answer this question, at least partially, it is helpful to decompose the angular part of the wave into spherical harmonics. Aretakis [2011a] proved indeed an (I)LED statement with loss of one derivative for waves that are supported on the angular frequencies $l \ge 1$. By constructing a localised solution with vanishing spherical mean we can show that this result is optimal in the sense that *some* loss of derivative is again necessary. This can be done for instance by considering the superposition of two Gaussian beams that follow the generators $\gamma_1(s) = \big(s, m, \frac{1}{2}\pi, \frac{1}{2}\pi\big)$ and $\gamma_2(s) = \big(s, m, \frac{1}{2}\pi, \frac{3}{2}\pi\big)$, where the initial value of beam one is exactly the negative of the initial value of beam two if translated in the φ variable by π.[38] The question of whether one can prove an ILED statement with loss of derivative in the case $l = 0$ is still open, though it is expected that the answer is negative. In order to obtain stability results for waves supported on all angular frequencies, Aretakis had to use the degenerate energy (of course these results are weaker than results one would obtain if an ILED statement for the case $l = 0$ actually held).

7. Applications to Kerr black holes

The Kerr family is a 2-parameter family of solutions to the vacuum Einstein equations. Let us fix the manifold $M := \mathbb{R} \times (m + \sqrt{m^2 - a^2}, \infty) \times \mathbb{S}^2$, where m and a are real parameters that will model the mass and the angular momentum per unit mass of the black hole, respectively, and which are restricted to the range $0 \le a \le m$, $0 \ne m$. Let (t, r, θ, φ) denote the standard coordinates on the manifold M and define functions

$$\rho^2 := r^2 + a^2\cos^2\theta, \qquad g_{tt} := -1 + \frac{2mr}{\rho^2},$$
$$\Delta := r^2 - 2mr + a^2, \qquad g_{t\varphi} := -\frac{2mra\sin^2\theta}{\rho^2},$$
$$g_{\varphi\varphi} := \left(r^2 + a^2 + \frac{2mra^2\sin^2\theta}{\rho^2}\right)\sin^2\theta.$$

[38]Let us mention here that in this particular situation the approximation using geometric optics is easier. Indeed, one can easily write down a solution of the eikonal equation such that the characteristics are the outgoing null geodesics. First one has to prove then the analogue of Theorem 4.1, which is easier since the approximate conservation law we used in the case of Gaussian beams is replaced by an exact conservation law for the geometric optics approximation; see footnote 23 on page 1396. But then we can easily contradict the validity of (I)LED statements for any angular frequency: working in $(t^*, r, \theta, \varphi)$ coordinates, we choose the initial value of the function a in the transport equation (i.e., the second equation in (1.6)) to have the angular dependence of a certain spherical harmonic and the radial dependence corresponds to a smooth cut-off, i.e., a initially is only nonvanishing for $r \in [m, m + \varepsilon)$.

The metric on M is then defined by

$$g = g_{tt}\, dt^2 - g_{t\varphi}(d\varphi \otimes dt + dt \otimes d\varphi) + g_{\varphi\varphi}\, d\varphi^2 + \frac{\rho^2}{\Delta}\, dr^2 + \rho^2\, d\theta^2.$$

The roots of $\Delta(r)$ are denoted by r_- and r_+, where $r_\pm = m \pm \sqrt{m^2 - a^2}$. As for the Reissner–Nordström family, one can (and should) extend these spacetimes in order to understand their physical interpretation as a black hole. For details, we refer the reader again to [Hawking and Ellis 1973]. Fixing the θ coordinate to be $\frac{1}{2}\pi$ and modding out the $\mathbb{S}^1$ corresponding to the φ coordinate, we again obtain pictorial representations of these spacetimes. For the subextremal case $0 < a < m$, the diagram is the same as the one depicted in Section 6C, while, in the extremal case $a = m$, one obtains the same diagram as in Section 6F.

7A. ***Trapping in (sub)extremal Kerr.*** As in the case of the Schwarzschild spacetime there are trapped null geodesics in the domain of outer communications of the Kerr spacetime whose energy stays bounded away from zero and infinity if the energy-measuring vector field N is sensibly chosen. In the case of $a > 0$, however, the set that accommodates trapped null geodesics is the closure of an *open* set in spacetime, which is in contrast to the 3-dimensional photonsphere in Schwarzschild and Reissner–Nordström. Before we explain in some more detail how to find the trapped geodesics, we set up a suitable choice of foliation and energy-measuring vector field:

For (sub)extremal Kerr we foliate the domain of outer communication (which is covered by the above (t, r, θ, φ) coordinates) in the same way as we did before for the Schwarzschild and the extremal Reissner–Nordström spacetimes, namely by first introducing an ingoing "null" coordinate v and then subtracting off r to get a good time coordinate t^*. Slightly more general than is needed at this point, let us define

$$v_+ := t + r^* \quad \text{and} \quad \varphi_+ := \varphi + \bar{r},$$

where r^* is defined up to a constant by $dr^*/dr = (r^2 + a^2)/\Delta$ and $\bar{r}$ is defined up to a constant by $d\bar{r}/dr = a/\Delta$. The set of functions $(v_+, r, \theta, \varphi_+)$ form ingoing "null" coordinates (v_+ is here the "null" coordinate, however, it does *not* satisfy the eikonal equation $d\phi \cdot d\phi = 0$), they cover the regions *I*, *II* and *III* in the spacetime diagram for subextremal Kerr,[39] and the metric takes the form

$$\begin{aligned} g = g_{tt}\, dv_+^2 + g_{t\varphi}(dv_+ \otimes d\varphi_+ + d\varphi_+ \otimes dv_+) + (dv_+ \otimes dr + dr \otimes dv_+) \\ - a \sin^2\theta (dr \otimes d\varphi_+ + d\varphi_+ \otimes dr) + g_{\varphi\varphi}\, d\varphi_+^2 + \rho^2\, d\theta^2. \end{aligned}$$

Finally, we define $t^* := v_+ - r$. That this is indeed a good time coordinate is easily seen from writing the metric in $(t^*, r, \theta, \varphi_+)$ coordinates and restricting it to $\{t^* = \text{const}\}$ slices: One obtains

$$\bar{g} = (g_{tt} + 2)\, dr^2 + (g_{t\varphi} - a \sin^2\theta)(d\varphi_+ \otimes dr + dr \otimes d\varphi_+) + \rho^2\, d\theta^2 + g_{\varphi\varphi}\, d\varphi_+^2,$$

and the (θ, θ) minor of this matrix is found to be $2mr \sin^2\theta + (r^2 + a^2)\sin^2\theta - a^2 \sin^4\theta$, which is positive away from the well-understood coordinate singularity $\theta = \{0, \frac{1}{2}\pi\}$. Hence, the slices $\Sigma_\tau := \{t^* = \tau\}$ are spacelike and it is easily seen that they asymptote to $\{t = \text{const}\}$ slices near spacelike infinity and end

[39] In the extremal case they cover all of the spacetime diagram depicted in Figure 8 in Section 6F.

on the *future* event horizon. A suitable timelike vector field N for measuring the energy is again given by $N := -(dt^*)^\sharp$.

To be more precise about what we mean by a null geodesic being trapped, let us call a future, complete, affinely parametrised null geodesic $\gamma : [0, \infty) \to M$ (which is in particular contained in the black hole exterior M) *trapped* if, and only if, it does not escape to infinity, i.e., for $s \to \infty$ we do *not* have $(r \circ \gamma)(s) \to \infty$. In the following we give a brief sketch of how one finds the trapped null geodesics. For a detailed discussion of the geodesic flow we refer the reader to [O'Neill 1995] or [Chandrasekhar 1998].

The starting point for the investigation of the behaviour of the geodesics in the Kerr spacetime is the observation that the geodesic flow separates. An affinely parametrised *null* geodesic $\gamma(s) = (t(s), r(s), \theta(s), \varphi(s))$ satisfies the following first-order equations:

$$\rho^2 \dot{t} = a\mathbb{D} + (r^2 + a^2)\frac{\mathbb{P}}{\Delta}, \tag{7.1}$$

$$\rho^4 (\dot{r})^2 = R(r) := -\mathcal{K}\Delta + \mathbb{P}^2, \tag{7.2}$$

$$\rho^4 (\dot{\theta})^2 = \Theta(\theta) := \mathcal{K} - \frac{\mathbb{D}^2}{\sin^2\theta}, \tag{7.3}$$

$$\rho^2 \dot{\varphi} = \frac{\mathbb{D}}{\sin^2\theta} + \frac{a\mathbb{P}}{\Delta},$$

where $\mathcal{K}$ is the Carter constant of the geodesic, $\mathbb{P}(r) = (r^2 + a^2)E - La$ and $\mathbb{D}(\theta) = L - Ea\sin^2\theta$. Here, $E = -g(\partial_t, \dot{\gamma})$ is the energy of the geodesic[40] and $L = g(\partial_\varphi, \dot{\gamma})$ is the angular momentum. Note that since the left-hand side of (7.3) is positive, it follows that the Carter constant $\mathcal{K}$ is nonnegative.

In order to find all trapped null geodesics, the investigation naturally starts with (7.2). The crucial observation is that a simple zero of $R(r)$ corresponds to a turning point (in the r-coordinate) of the geodesic, while a double zero corresponds to an orbit of constant r (or to asymptotic approach).[41] It follows that a necessary condition for a null geodesic being trapped is that the constants of motion $\mathcal{K}$, L, and E can be chosen in such a way that either $R(r)$ has a double zero in (r_+, ∞) or $R(r)$ has two simple zeros in (r_+, ∞) and is nonnegative in between. In the following we show that the latter case cannot occur.

We distinguish the two cases $E = 0$ and $E \neq 0$. In the first case, $R(r)$ is a polynomial of order two with $R(r) \to -\infty$ for $r \to \infty$ (recall that $\mathcal{K} \geq 0$). Moreover, $R(r)$ is nonnegative in $[r_-, r_+]$. This shows that $R(r)$ can have at most one real root in (r_+, ∞).

In the case $E \neq 0$, $R(r)$ is a polynomial of order four. Over the complex numbers, we can write $R(r)$ as

$$R(r) = E^2 \cdot (r - \lambda_1)(r - \lambda_2)(r - \lambda_3)(r - \lambda_4) = E^2 \cdot r^4 - E^2(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \cdot r^3 + \cdots,$$

where $\lambda_i \in \mathbb{C}$, $i \in \{1, 2, 3, 4\}$, are the complex roots of $R(r)$. Since $R(r)$ does not have a term of order three, we see that the sum of the complex roots of $R(r)$ has to yield zero. This directly excludes $R(r)$

[40] Note that ∂_t is not timelike everywhere! However, one still calls this quantity the "energy" of the null geodesic.

[41] See Proposition 4.3.7 and Corollary 4.3.8 in Chapter 4 of [O'Neill 1995]

having four positive zeros. We also note that $R(r)$ tends to ∞ for $r \to \infty$; hence, for $R(r)$ to have two simple zeros in (r_+, ∞) and to be nonnegative in between, we see that $R(r)$ has to have at least three zeros in (r_+, ∞). But since $\mathcal{K} \geq 0$, we see that $R(r)$ is nonnegative in $[r_-, r_+]$; i.e., if $R(r)$ has three zeros in (r_+, ∞), then it needs to have a fourth positive zero, which we have already ruled out. This shows that trapping can only occur due to a double zero of $R(r)$.

We now sketch how one finds the values of r that accommodate trapped null geodesics (along with the constants of motion $\mathcal{K}$, L and E). A detailed discussion is found in Section 63(c) of [Chandrasekhar 1998].

Without loss of generality we can assume that $E = 1$. We then need to solve

$$R(r) = -\mathcal{K}(r^2 - 2mr + a^2) + (r^2 + a^2 - La)^2 = 0,$$
$$\frac{d}{dr}R(r) = 2\mathcal{K}(m - r) + 4r(r^2 + a^2 - La) = 0.$$

Eliminating $\mathcal{K}$, we obtain the two solutions

$$L_1(r) = \frac{r^2 + a^2}{a} \quad \text{and} \quad L_2(r) = \frac{r^3 + ra^2 - 3mr^2 + ma^2}{a(m - r)}.$$

In the first case we obtain $\mathcal{K}_1(r) = 0$, which characterises the principal null geodesics (see Corollary 4.2.8 in [O'Neill 1995]) and is thus not compatible with orbits of constant r. We are thus left with the second solution $L_2(r)$, which implies $\mathcal{K}_2(r) = (4r^2/(m-r)^2)\Delta$. For the further analysis it is helpful to introduce the quantity $\mathcal{Q} = \mathcal{K} - (L - a)^2$, since it simplifies the analysis of the θ-motion of the geodesic. We obtain

$$\mathcal{Q}_2(r) = \frac{r^3}{a^2(m - r)^2}(4a^2m - r(r - 3m)^2).$$

It can now be shown (see Section 63(c) of [Chandrasekhar 1998]) that if we evaluate the right-hand side of (7.3) at $L_2(r)$ and $\mathcal{K}_2(r)$, where r is such that $\mathcal{Q}_2(r) < 0$, then we see that it is negative for all values of θ. Hence, these values of r do not accommodate trapped null geodesics. However, one can show that the values of r where $\mathcal{Q}_2(r) \geq 0$ indeed allow the presence of trapped null geodesics. This region is bounded by the roots r_δ and r_ρ of $\mathcal{Q}_2(r)$, which are bigger than r_+.

We now show that the N-energy of a trapped null geodesic γ_{r_0}, trapped on the hypersurface $\{r = r_0\}$ with $r_0 \in [r_\delta, r_\rho]$, is bounded away from zero and infinity. One way to do this is to compute the N-energy directly:

$$-(N, \dot{\gamma}) = (dt + dr^* - dr)(\dot{\gamma}) = \dot{t} = \frac{1}{\rho^2}\left[a\mathbb{D}(\theta) + (r_0^2 + a^2)\frac{\mathbb{P}(r_0)}{\Delta(r_0)}\right]$$

where we have used (7.1). A further analysis of the behaviour of the θ component of γ_{r_0} yields that its image is a closed subset of $[0, \pi]$; thus $-(N, \dot{\gamma})(\theta)$ takes on its minimum and maximum. Since $-(N, \dot{\gamma})$ is always strictly positive, this immediately yields that it is bounded away from zero and infinity.

Invoking Theorem 5.5 we thus obtain:

Theorem 7.4 (trapping in (sub)extremal Kerr). *Let (M, g) be the domain of outer communications of a (sub)extremal Kerr spacetime, foliated by the level sets of a time function t^* as above. Moreover, let N*

be the timelike vector field from above and $\mathcal{T}$ an open set with the property that for all $\tau \geq 0$ we have $\mathcal{T} \cap \Sigma_\tau \cap [r_\delta, r_\rho] \neq \varnothing$. Then there is no function $P : [0, \infty) \to (0, \infty)$ with $P(\tau) \to 0$ for $\tau \to \infty$ such that

$$E^N_{\tau, \mathcal{T}\cap\Sigma_\tau}(u) \leq P(\tau) E^N_0(u)$$

holds for all solutions u of the wave equation.

Note that the same remark as made in footnote 27 on page 1400 applies: the theorem remains true if we choose a different timelike vector field N which commutes with the Killing vector field ∂_t and also if we choose a different foliation by timelike slices, i.e., a different time function.[42]

Another way to show that the energy of the trapped null geodesic γ_{r_0} is bounded away from zero and infinity is to choose a different suitable vector field N. Recall that the vector fields ∂_t and ∂_φ are Killing, and that at each point in the domain of outer communications they also span a timelike direction. We can thus find a timelike vector field $\tilde{N}$ that commutes with ∂_t and such that in a small r-neighbourhood of r_0 the vector field $\tilde{N}$ is given by $\partial_t + k\partial_\varphi$ with $k \in \mathbb{R}$ a constant. Thus, $\tilde{N}$ is Killing in this small r-neighbourhood and, hence, the $\tilde{N}$-energy of γ_{r_0} is constant.

7B. *Blue-shift near the Cauchy horizon of (sub)extremal Kerr.* In this section we show that the results of Section 6C and 6D also hold for (sub)extremal Kerr. The proof is completely analogous: In the above defined $(v_+, r, \theta, \varphi_+)$ coordinates a family of ingoing null geodesics with uniformly bounded energy on Σ_0 near spacelike infinity ι^0 is given by

$$\gamma_{v_+^0}(s) = (v_+^0, -s, \theta_0, \varphi_0),$$

where $s \in (-\infty, 0)$. The same pictures as in Sections 6C and 6D apply, along with the same spacelike hypersurfaces Σ_0 and Σ_1. In order to obtain regular coordinates in a neighbourhood of the Cauchy horizon, we define, starting with (t, r, θ, φ) coordinates in region *II*, outgoing "null" coordinates $(v_-, r, \theta, \varphi_-)$ by $v_- = t - r^*$ and $\varphi_- = \varphi - \bar{r}$. These coordinates cover the regions *II* and *IV* in the subextremal case and regions *II* and *III* in the extremal case. In these coordinates, the tangent vector of the null geodesic $\gamma_{v_+^0}$ takes the form

$$\dot{\gamma}_{v_+^0} = -\frac{\partial}{\partial r}\bigg|_+ = 2\frac{r^2+a^2}{\Delta}\frac{\partial}{\partial v_-}\bigg|_- - \frac{\partial}{\partial r}\bigg|_- + 2\frac{a}{\Delta}\frac{\partial}{\partial \varphi_-}\bigg|_-, \tag{7.5}$$

which blows up at the Cauchy horizon. It is again easy to see that the inner product with a timelike vector field, which extends smoothly to a timelike vector field over the Cauchy horizon, necessarily blows up along Σ_1 for $v_+^0 \to \infty$. Thus, we obtain, after invoking Theorem 5.1:

Theorem 7.6 (blue-shift near the Cauchy horizon in subextremal Kerr). *Let Σ_0 and Σ_1 be spacelike slices in the subextremal Kerr spacetime as indicated in Figure 5 in Section 6C. Then there exists a sequence $\{u_i\}_{i\in\mathbb{N}}$ of solutions to the wave equation with initial energy $E_0^{n_{\Sigma_0}}(u_i) = 1$ on Σ_0 such that the n_{Σ_1}-energy on Σ_1 goes to infinity, i.e., $E_1^{n_{\Sigma_1}}(u_i) \to \infty$ for $i \to \infty$.*

In particular, there is no energy boundedness statement of the form (6.8).

[42] In the latter case one may have to alter the decay statement for the function P, i.e., replace it with $P(\tau) \to 0$ for $\tau \to \tau^*$.

As before, let us state the following:

Conjecture 7.7. *For generic compactly supported smooth initial data on Σ_0, the n_{Σ_1}-energy along Σ_1 of the corresponding solution to the wave equation is infinite.*

Let us conclude this section with a few remarks:

(i) Obviously, an analogous statement to Theorem 7.6 is true for extremal Kerr, however, one has to introduce again a suitable globally hyperbolic subset in order to be able to apply Theorem 5.1.

(ii) The discussion in Section 6E carries over to the Kerr case. In particular let us stress that Conjecture 7.7 only concerns *sub*extremal Kerr black holes — the same statement for extremal Kerr black holes is expected to be false. However, as for Reissner–Nordström black holes, we conjecture a C^k instability (for some finite k) at the Cauchy horizon of extremal Kerr black holes.

(iii) We leave it as an exercise for the reader to convince him- or herself that analogous versions of Theorems 7.4 and 7.6 also hold true for the Kerr–Newman family.

Appendix: A breakdown criterion for solutions of the eikonal equation

We give a breakdown criterion for solutions of the eikonal equation for which a given null geodesic is a characteristic.

Theorem A.1. *Let (M, g) be a Lorentzian manifold and $\gamma : [0, a) \to M$ an affinely parametrised null geodesic, $a \in (0, \infty]$. If γ has conjugate points then there exists no solution $\phi : U \to \mathbb{R}$ of the eikonal equation $d\phi \cdot d\phi = 0$ with $\operatorname{grad}\phi\big|_{\operatorname{Im}\gamma} = \dot{\gamma}$, where U is a neighbourhood of $\operatorname{Im}\gamma$.*

The theorem is motivated by the construction of localised solutions to the wave equation using the naive geometric optics approximation, where we need to find a solution of the eikonal equation for which a given null geodesic is a characteristic; see (1.6). It is well known that solutions of the eikonal equation break down whenever characteristics cross. However, by choosing the initial data (and thus the neighbouring characteristics) suitably one can try to avoid crossing characteristics. This is for example possible in the Minkowski spacetime. The theorem gives a sufficient condition for when no such choice is possible.

Our proof is a minor adaptation of Riemannian methods to the Lorentzian null case; see, for example, [Eschenburg and O'Sullivan 1976], in particular their Proposition 3.

First we need some groundwork. We pull back the tangent bundle TM via γ and denote the subbundle of vectors that are orthogonal to $\dot{\gamma}$ by $N(\gamma)$. The vectors that are proportional to $\dot{\gamma}$ give rise to a subbundle of $N(\gamma)$, which we quotient out to obtain the quotient bundle $\overline{N}(\gamma)$. It is easy to see that the metric g induces a positive-definite metric $\bar{g}$ on $\overline{N}(\gamma)$ and that the bundle map $R_\gamma : N(\gamma) \to N(\gamma)$, where $R_\gamma(X) = R(X, \dot{\gamma})\dot{\gamma}$ and R is the Riemann curvature tensor, induces a bundle map $\bar{R}_\gamma$ on $\overline{N}(\gamma)$, and finally that the Levi-Civita connection ∇ induces a connection $\overline{\nabla}$ for $\overline{N}(\gamma)$.

Definition A.2. $\bar{J} \in \operatorname{End}(\overline{N}(\gamma))$ is a *Jacobi tensor class* if and only if $\overline{D}_t^2 \bar{J} + \bar{R}_\gamma \bar{J} = 0$.[43]

[43] Here and in what follows we write $\overline{D}_t$ for $\overline{\nabla}_{\partial_t}$.

A Jacobi tensor class should be thought of as a variation field of γ that arises from a many-parameter variation by geodesics. It generalises the notion of a Jacobi field (class), an infinitesimal 1-parameter variation. Indeed, a solution ϕ of the eikonal equation for which γ is a characteristic gives rise to a Jacobi tensor class $\bar{J}$:

We denote the flow of $\operatorname{grad}\phi$ by Ψ_t and define $J \in \operatorname{End}(N(\gamma))$ by

$$J_t(X_t) := (\Psi_t)_*(X_0),$$

where we extend $X_t \in N(\gamma)_t$ by parallel propagation to a vector field X along γ whose value at 0 is X_0. Note that J is well defined, that is, we have $J_t(X_t) \in N(\gamma)$: Given $X_0 \in T_{\gamma(0)}M$, extend it to a vector field $\tilde{X}$ on M with $[\tilde{X}, \operatorname{grad}\phi] = 0$, i.e., along γ we have $\tilde{X}\big|_{\gamma(t)} = (\Psi_t)_*(X_0)$. Then

$$0 = \nabla_{\tilde{X}}(\operatorname{grad}\phi, \operatorname{grad}\phi) = 2(\nabla_{\tilde{X}} \operatorname{grad}\phi, \operatorname{grad}\phi) = 2\nabla_{\operatorname{grad}\phi}(\tilde{X}, \operatorname{grad}\phi),$$

from which it follows that $\tilde{X}\big|_{\gamma(t)}$ is orthogonal to $\operatorname{grad}\phi\big|_{\gamma(t)}$. Moreover, J is a Jacobi tensor:[44] Let X be a parallel section along γ and $\tilde{X}$ an extension of X_0 as above. Then

$$(D_t J)(X) = D_t(JX) = D_t(\Psi_{t*}X_0) = \nabla_{\operatorname{grad}\phi}\tilde{X} = \nabla_{\tilde{X}} \operatorname{grad}\phi = \nabla_{JX} \operatorname{grad}\phi.$$

Thus,

$$D_t J = (\nabla \operatorname{grad}\phi) \circ J. \tag{A.3}$$

Differentiating once more gives

$$(D_t^2 J)(X) = \nabla_{\operatorname{grad}\phi}(\nabla_{JX} \operatorname{grad}\phi) = R(\operatorname{grad}\phi, JX) \operatorname{grad}\phi = -R_\gamma \circ J(X).$$

Using that $(\Psi_t)_*(\operatorname{grad}\phi|_{\gamma(0)}) = \operatorname{grad}\phi|_{\gamma(t)}$, it is now clear that J descends to a Jacobi tensor class $\bar{J}$. Moreover, $\bar{J}$ is nonsingular, i.e., $\bar{J}^{-1}$ exists. Since the metric $\bar{g}$ is nondegenerate, we can form adjoints of sections of $\operatorname{End}(\overline{N}(\gamma))$, which we will denote by *. Note also that $(\overline{D}_t\bar{J})\bar{J}^{-1}$ is self-adjoint. This follows from (A.3) and the fact that $\nabla\nabla\phi$ is symmetric. We now prove the theorem.

Proof of Theorem A.1. Assume there exists such a solution ϕ of the eikonal equation. Say the points $\gamma(t_0)$ and $\gamma(t_1)$ are conjugate, $0 \le t_0 < t_1 < a$, and $\bar{J}$ is the Jacobi tensor class induced by ϕ, as discussed above. Using the identification of $\operatorname{End}(\overline{N}(\gamma)_t)$ with $\operatorname{End}(\overline{N}(\gamma)_{t_0})$ via parallel translation, we write

$$\overline{K}(t) := \bar{J}(t)C\int_{t_0}^{t} (\bar{J}^*\bar{J})^{-1}(\tau)\,d\tau,$$

where $C = \bar{J}^{-1}(t_0)\bar{J}^*(t_0)\bar{J}(t_0)$. A straightforward computation shows that $\overline{K}$ is a Jacobi tensor class with $\overline{K}(t_0) = 0$ and $\overline{D}_t\overline{K}(t_0) = \mathrm{id}$. Moreover, $\overline{K}(t)$ is nonsingular for $t > t_0$.

On the other hand, there exists a Jacobi field Y with $Y(t_0) = 0$ and $Y(t_1) = 0$. This implies that Y is a section of $N(\gamma)$. The Jacobi field Y induces a nontrivial Jacobi field class $\bar{Y}$ that vanishes at t_0 and t_1. However, a Jacobi field class is uniquely determined by its value and velocity at a point. Parallelly propagating $\overline{D}_t\bar{Y}\big|_{t_0}$ gives rise to a vector field class $\bar{Z}$. Then $\overline{K}\bar{Z}$ is a Jacobi field class that has the same value and velocity as $\bar{Y}$ at $t = t_0$, thus $\overline{K}\bar{Z} = \bar{Y}$. This, however, contradicts $\overline{K}$ being nonsingular for $t > t_0$. $\square$

[44] This notion is analogous to Definition A.2, without taking the quotient.

Acknowledgements

I would like to thank my supervisor Mihalis Dafermos for numerous instructive and stimulating discussions. Moreover, I am grateful to Mihalis Dafermos, Gustav Holzegel, Stefanos Aretakis, and two anonymous referees for many useful comments on a preliminary version of this paper. Furthermore, I would like to thank the Science and Technology Facilities Council (STFC), the European Research Council (ERC), and the German Academic Exchange Service (DAAD) (Doktorandenstipendium) for their financial support.

References

[Andersson and Blue 2009] L. Andersson and P. Blue, "Hidden symmetries and decay for the wave equation on the Kerr spacetime", preprint, 2009. arXiv 0908.2265v2

[Aretakis 2011a] S. Aretakis, "Stability and instability of extreme Reissner–Nordström black hole spacetimes for linear scalar perturbations, I", *Comm. Math. Phys.* **307**:1 (2011), 17–63. MR 2835872 Zbl 1229.85002

[Aretakis 2011b] S. Aretakis, "Stability and instability of extreme Reissner-Nordström black hole spacetimes for linear scalar perturbations, II", *Ann. Henri Poincaré* **12**:8 (2011), 1491–1538. MR 2855176 Zbl 1242.83049

[Arnaud 1973] J. A. Arnaud, "Hamiltonian theory of beam mode propagation", pp. 249–304 in *Progress in Optics*, vol. XI, edited by E. Wolf, North-Holland, Amsterdam, 1973.

[Babich and Buldyrev 2009] V. M. Babich and V. S. Buldyrev, *Asymptotic methods in short-wavelength diffraction theory*, Alpha Science, Oxford, 2009. MR 54 #14569 Zbl 0742.35002

[Babich and Ulin 1981] V. M. Babich and V. V. Ulin, "The complex space-time ray method and "quasiphotons"", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **117** (1981), 5–12. In Russian; translated in *J. Soviet Math.* **24**:3 (1984), 269–273. MR 83j:35094 Zbl 0477.35025

[Bernal and Sánchez 2005] A. N. Bernal and M. Sánchez, "Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes", *Comm. Math. Phys.* **257**:1 (2005), 43–50. MR 2006g:53105 Zbl 1081.53059

[Blue and Soffer 2009] P. Blue and A. Soffer, "Phase space analysis on some black hole manifolds", *J. Funct. Anal.* **256**:1 (2009), 1–90. MR 2010i:83072 Zbl 1158.83007

[Blue and Sterbenz 2006] P. Blue and J. Sterbenz, "Uniform decay of local energy and the semi-linear wave equation on Schwarzschild space", *Comm. Math. Phys.* **268**:2 (2006), 481–504. MR 2007i:58037 Zbl 1123.58018

[Chandrasekhar 1998] S. Chandrasekhar, *The mathematical theory of black holes*, Clarendon/Oxford University Press, New York, 1998. MR 99f:83041 Zbl 0912.53053

[Choquet-Bruhat 2009] Y. Choquet-Bruhat, *General relativity and the Einstein equations*, Oxford University Press, Oxford, 2009. MR 2010f:83001 Zbl 1157.83002

[Civin 2014] D. Civin, *Stability of charged rotating black holes for linear scalar perturbations*, Ph.D. thesis, University of Cambrdige, 2014, available at https://www.repository.cam.ac.uk/handle/1810/247397.

[Dafermos 2005] M. Dafermos, "The interior of charged black holes and the problem of uniqueness in general relativity", *Comm. Pure Appl. Math.* **58**:4 (2005), 445–504. MR 2006e:83087 Zbl 1071.83037

[Dafermos and Rodnianski 2009] M. Dafermos and I. Rodnianski, "The red-shift effect and radiation decay on black hole spacetimes", *Comm. Pure Appl. Math.* **62**:7 (2009), 859–919. MR 2011b:83059 Zbl 1169.83008

[Dafermos and Rodnianski 2010a] M. Dafermos and I. Rodnianski, "Decay for solutions of the wave equation on Kerr exterior spacetimes I–II: The cases $|a| \ll M$ or axisymmetry", preprint, 2010. arXiv 1010.5132v1

[Dafermos and Rodnianski 2010b] M. Dafermos and I. Rodnianski, "A new physical-space approach to decay for the wave equation with applications to black hole spacetimes", pp. 421–432 in *XVIth International Congress on Mathematical Physics*, edited by P. Exner, World Science, Hackensack, NJ, 2010. MR 2012e:58051 Zbl 1211.83019

[Dafermos and Rodnianski 2011a] M. Dafermos and I. Rodnianski, "The black hole stability problem for linear scalar perturbations", pp. 132–189 in *Proceedings of the Twelfth Marcel Grossman Meeting on General Relativity* (Paris, 2009), edited by T. Damour et al., World Scientific, Singapore, 2011.

[Dafermos and Rodnianski 2011b] M. Dafermos and I. Rodnianski, "A proof of the uniform boundedness of solutions to the wave equation on slowly rotating Kerr backgrounds", *Invent. Math.* **185**:3 (2011), 467–559. MR 2827094 Zbl 1226.83029

[Dafermos and Rodnianski 2013] M. Dafermos and I. Rodnianski, "Lectures on black holes and linear waves", pp. 97–205 in *Evolution equations*, edited by D. Ellwood et al., Clay Math. Proc. **17**, Amer. Math. Soc., Providence, RI, 2013. MR 3098640 Zbl 1300.83004

[Dafermos et al. 2013] M. Dafermos, G. Holzegel, and I. Rodnianski, "A scattering theory construction of dynamical vacuum black holes", preprint, 2013. arXiv 1306.5364v2

[Dyatlov 2011] S. Dyatlov, "Exponential energy decay for Kerr–de Sitter black holes beyond event horizons", *Math. Res. Lett.* **18**:5 (2011), 1023–1035. MR 2875874 Zbl 1253.83020

[Eschenburg and O'Sullivan 1976] J.-H. Eschenburg and J. J. O'Sullivan, "Growth of Jacobi fields and divergence of geodesics", *Math. Z.* **150**:3 (1976), 221–237. MR 54 #6194 Zbl 0318.53044

[Hawking and Ellis 1973] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, Cambridge Monographs on Mathematical Physics **1**, Cambridge University Press, 1973. MR 54 #12154 Zbl 0265.53054

[Holzegel and Smulevici 2013] G. Holzegel and J. Smulevici, "Decay properties of Klein–Gordon fields on Kerr–AdS spacetimes", *Comm. Pure Appl. Math.* **66**:11 (2013), 1751–1802. MR 3103910 Zbl 1277.83023

[Keller 1956] J. B. Keller, "Diffraction by a convex cylinder", *Trans. IRE* **4**:3 (1956), 312–321. MR 20 #641 Zbl 1206.22004

[Klainerman 1985] S. Klainerman, "Uniform decay estimates and the Lorentz invariance of the classical wave equation", *Comm. Pure Appl. Math.* **38**:3 (1985), 321–332. MR 86i:35091 Zbl 0635.35059

[Klainerman et al. 2013] S. Klainerman, I. Rodnianski, and J. Szeftel, "The bounded L^2 curvature conjecture", preprint, 2013. arXiv 1204.1767v2

[Lucietti and Reall 2012] J. Lucietti and H. S. Reall, "Gravitational instability of an extreme Kerr black hole", *Phys. Rev. D* **86** (2012), Art. ID #104030.

[Luk 2010] J. Luk, "Improved decay for solutions to the linear wave equation on a Schwarzschild black hole", *Ann. Henri Poincaré* **11**:5 (2010), 805–880. MR 2012k:58045 Zbl 1208.83068

[Maslov 1994] V. P. Maslov, *The complex WKB method for nonlinear equations, I: Linear theory*, Progress in Physics **16**, Birkhäuser, Basel, 1994. MR 95m:58126 Zbl 0811.35088

[Metcalfe et al. 2012] J. Metcalfe, D. Tataru, and M. Tohaneanu, "Price's law on nonstationary space-times", *Adv. Math.* **230**:3 (2012), 995–1028. MR 2921169 Zbl 1246.83070

[Misner et al. 1973] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Francisco, 1973. MR 54 #6869

[Morawetz 1961] C. S. Morawetz, "The decay of solutions of the exterior initial-boundary value problem for the wave equation", *Comm. Pure Appl. Math.* **14** (1961), 561–568. MR 24 #A2744 Zbl 0101.07701

[Morawetz 1962] C. S. Morawetz, "The limiting amplitude principle", *Comm. Pure Appl. Math.* **15** (1962), 349–361. MR 27 #1696 Zbl 0196.41202

[Morawetz 1968] C. S. Morawetz, "Time decay for the nonlinear Klein–Gordon equations", *Proc. Roy. Soc. Ser. A* **306** (1968), 291–296. MR 38 #2455 Zbl 0157.41502

[Murata et al. 2013] K. Murata, H. S. Reall, and N. Tanahashi, "What happens at the horizon(s) of an extreme black hole?", *Classical Quantum Gravity* **30**:23 (2013), Art. ID #235007. MR 3129282 Zbl 1284.83093

[O'Neill 1983] B. O'Neill, *Semi-Riemannian geometry: with applications to relativity*, Pure and Applied Mathematics **103**, Academic Press, New York, 1983. MR 85f:53002 Zbl 0531.53051

[O'Neill 1995] B. O'Neill, *The geometry of Kerr black holes*, A K Peters, Wellesley, MA, 1995. MR 96c:83052 Zbl 0828.53078

[Penrose 1968] R. Penrose, "Structure of space-time", in *Battelle Rencontres* (Seattle, 1967), edited by C. DeWitt-Morette and J. A. Wheeler, W. A. Benjamin, New York, 1968. Zbl 0174.55901

[Ralston 1969] J. V. Ralston, "Solutions of the wave equation with localized energy", *Comm. Pure Appl. Math.* **22** (1969), 807–823. MR 40 #7642 Zbl 0209.40402

[Ralston 1976] J. V. Ralston, "On the construction of quasimodes associated with stable periodic orbits", *Comm. Math. Phys.* **51**:3 (1976), 219–242. MR 54 #14003 Zbl 0333.35066

[Ralston 1982] J. Ralston, "Gaussian beams and the propagation of singularities", pp. 206–248 in *Studies in partial differential equations*, edited by W. Littman, MAA Stud. Math. **23**, Math. Assoc. America, Washington, DC, 1982. MR 85c:35052 Zbl 0533.35062

[Schlue 2013] V. Schlue, "Decay of linear waves on higher-dimensional Schwarzschild black holes", *Anal. PDE* **6**:3 (2013), 515–600. MR 3080190 Zbl 06212808

[Tataru and Tohaneanu 2011] D. Tataru and M. Tohaneanu, "A local energy estimate on Kerr black hole backgrounds", *Int. Math. Res. Not.* **2011**:2 (2011), 248–292. MR 2012a:58050 Zbl 1209.83028

[Taylor 2011] M. E. Taylor, *Partial differential equations, I: Basic theory*, 2nd ed., Applied Mathematical Sciences **115**, Springer, New York, 2011. MR 2011m:35001 Zbl 1206.35002

[Zenginoglu and Galley 2012] A. Zenginoglu and C. R. Galley, "Caustic echoes from a Schwarzschild black hole", *Phys. Rev. D* **86** (2012), Art. ID #064030.

Received 7 May 2014. Revised 24 Sep 2014. Accepted 30 Apr 2015.

JAN SBIERSKI: jjs48@cam.ac.uk
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom

ANALYSIS AND PDE
Vol. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1421

msp

HEIGHT ESTIMATE AND SLICING FORMULAS IN THE HEISENBERG GROUP

ROBERTO MONTI AND DAVIDE VITTONE

We prove a height estimate (distance from the tangent hyperplane) for Λ-minimizers of the perimeter in the sub-Riemannian Heisenberg group. The estimate is in terms of a power of the excess (L^2-mean oscillation of the normal) and its proof is based on a new coarea formula for rectifiable sets in the Heisenberg group.

1. Introduction

We continue the research project started in [Monti and Vittone 2012; Monti 2014] on the regularity of H-perimeter minimizing boundaries in the Heisenberg group $\mathbb{H}^n$. Our goal is to prove the so-called *height estimate* for sets that are Λ*-minimizers* and have small *excess* inside suitable cylinders; see Theorem 1.3. The proof follows the scheme of the median choice for the measure of the boundary in certain half-cylinders together with a lower-dimensional isoperimetric inequality on slices. For minimizing currents in $\mathbb{R}^n$, the principal ideas of the argument go back to [Almgren 1968] and are carried over in [Federer 1969, Theorem 5.3.4]. The argument can be also found in the Appendix of [Schoen and Simon 1982] and, for Λ-minimizers of perimeter in $\mathbb{R}^n$, in [Maggi 2012, Section 22.2]. For minimizers of H-perimeter, the decay estimate of excess of Almgren and De Giorgi is still an open problem; see [Monti 2015].

Our main technical effort is the proof of a coarea formula (slicing formula) for intrinsic rectifiable sets; see Theorem 1.5. This formula is established in Section 2 and has a nontrivial character because the domain of integration and its slices need not be rectifiable in the standard sense. The relative isoperimetric inequalities that are used in the slices reduce to a single isoperimetric inequality in one slice that is relative to a family of varying domains with uniform isoperimetric constants. This uniformity can be established using the results on regular domains in Carnot groups of step 2 in [Monti and Morbidelli 2005] and the isoperimetric inequality in [Garofalo and Nhieu 1996]; see Section 3A.

The $(2n+1)$-dimensional Heisenberg group is the manifold $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$, $n \in \mathbb{N}$, endowed with the group product

$$(z,t) * (\zeta,\tau) = (z+\zeta, t+\tau+2\,\Im\langle z,\bar{\zeta}\rangle), \tag{1-1}$$

where $t, \tau \in \mathbb{R}$, $z, \zeta \in \mathbb{C}^n$ and $\langle z,\bar{\zeta}\rangle = z_1\bar{\zeta}_1 + \cdots + z_n\bar{\zeta}_n$. The Lie algebra of left-invariant vector fields in $\mathbb{H}^n$ is spanned by the vector fields

$$X_j = \frac{\partial}{\partial x_j} + 2y_j\frac{\partial}{\partial t}, \quad Y_j = \frac{\partial}{\partial y_j} - 2x_j\frac{\partial}{\partial t}, \quad \text{and} \quad T = \frac{\partial}{\partial t}, \tag{1-2}$$

MSC2010: 49Q05, 53C17.
Keywords: Heisenberg group, regularity of H-minimal surfaces, height estimate, slicing formula.

with $z_j = x_j + iy_j$ and $j = 1, \dots, n$. We denote by H the horizontal subbundle of $T\mathbb{H}^n$. Namely, for any $p = (z, t) \in \mathbb{H}^n$ we let

$$H_p = \operatorname{span}\{X_1(p), \dots, X_n(p), Y_1(p), \dots, Y_n(p)\}.$$

A horizontal section $\varphi \in C_c^1(\Omega; H)$, where $\Omega \subset \mathbb{H}^n$ is an open set, is a vector field of the form

$$\varphi = \sum_{j=1}^{n} \varphi_j X_j + \varphi_{n+j} Y_j,$$

where $\varphi_j \in C_c^1(\Omega)$, i.e., each coordinate φ_j is a continuously differentiable function with compact support contained in Ω.

Let g be the left-invariant Riemannian metric on $\mathbb{H}^n$ that makes orthonormal the vector fields $X_1, \dots X_n$, $Y_1, \dots, Y_n, T$ in (1-2). For tangent vectors $V, W \in T\mathbb{H}^n$, we let

$$\langle V, W\rangle_g = g(V, W) \quad \text{and} \quad |V|_g = g(V, V)^{1/2}.$$

The sup norm with respect to g of a horizontal section $\varphi \in C_c^1(\Omega; H)$ is

$$\|\varphi\|_g = \max_{p\in\Omega} |\varphi(p)|_g.$$

The Riemannian divergence of φ is

$$\operatorname{div}_g \varphi = \sum_{j=1}^{n} X_j\varphi_j + Y_j\varphi_{n+j}.$$

The metric g induces a volume form on $\mathbb{H}^n$ that is left-invariant. Also, the Lebesgue measure $\mathscr{L}^{2n+1}$ on $\mathbb{H}^n$ is left-invariant, and by the uniqueness of the Haar measure the volume induced by g is the Lebesgue measure $\mathscr{L}^{2n+1}$. In fact, the proportionality constant is 1.

The *H-perimeter* of an $\mathscr{L}^{2n+1}$-measurable set $E \subset \mathbb{H}^n$ in an open set $\Omega \subset \mathbb{H}^n$ is

$$\mu_E(\Omega) = \sup\left\{ \int_E \operatorname{div}_g \varphi \, d\mathscr{L}^{2n+1} : \varphi \in C_c^1(\Omega; H), \|\varphi\|_g \le 1 \right\}.$$

If $\mu_E(\Omega) < \infty$ we say that E has finite H-perimeter in Ω. If $\mu_E(A) < \infty$ for any open set $A \Subset \Omega$, we say that E has locally finite H-perimeter in Ω. In this case, the open sets mapping $A \mapsto \mu_E(A)$ extends to a Radon measure μ_E on Ω that is called the *H-perimeter measure* induced by E. Moreover, there exists a μ_E-measurable function $\nu_E : \Omega \to H$ such that $|\nu_E|_g = 1$ μ_E-a.e. and the Gauss–Green integration by parts formula

$$\int_\Omega \langle \varphi, \nu_E\rangle_g \, d\mu_E = -\int_\Omega \operatorname{div}_g \varphi \, d\mathscr{L}^{2n+1}$$

holds for any $\varphi \in C_c^1(\Omega; H)$. The vector ν_E is called the *horizontal inner normal* of E in Ω.

The Korànyi norm of $p = (z, t) \in \mathbb{H}^n$ is $\|p\|_K = (|z|^4 + t^2)^{1/4}$. For any $r > 0$ and $p \in \mathbb{H}^n$, we define the balls

$$B_r = \{q \in \mathbb{H}^n : \|q\|_K < r\} \quad \text{and} \quad B_r(p) = \{p * q \in \mathbb{H}^n : q \in B_r\}.$$

The *measure-theoretic boundary* of a measurable set $E \subset \mathbb{H}^n$ is the set

$$\partial E = \big\{p \in \mathbb{H}^n : \mathscr{L}^{2n+1}(E \cap B_r(p)) > 0 \text{ and } \mathscr{L}^{2n+1}(B_r(p) \setminus E) > 0 \text{ for all } r > 0\big\}.$$

For a set E with locally finite H-perimeter, the H-perimeter measure μ_E is concentrated on ∂E and, actually, on a subset $\partial^* E$ of ∂E; see below. Moreover, up to modifying E on a Lebesgue-negligible set, one can always assume that ∂E coincides with the topological boundary of E; see [Serra Cassano and Vittone 2014, Proposition 2.5].

Definition 1.1. Let $\Omega \subset \mathbb{H}^n$ be an open set, $\Lambda \in [0, \infty)$, and $r \in (0, \infty]$. We say that a set $E \subset \mathbb{H}^n$ with locally finite H-perimeter in Ω is a (Λ, r)*-minimizer of H-perimeter in* Ω if, for any measurable set $F \subset \mathbb{H}^n$, $p \in \Omega$, and $s < r$ such that $E \Delta F \Subset B_s(p) \Subset \Omega$,

$$\mu_E(B_s(p)) \le \mu_F(B_s(p)) + \Lambda \mathscr{L}^{2n+1}(E \Delta F),$$

where $E \Delta F = E \setminus F \cup F \setminus E$.

We say that E is *locally H-perimeter minimizing in* Ω if, for any measurable set $F \subset \mathbb{H}^n$ and any open set U such that $E \Delta F \Subset U \Subset \Omega$, there holds $\mu_E(U) \le \mu_F(U)$.

We will often use the term Λ*-minimizer*, rather than (Λ, r)-minimizer, when the role of r is not relevant. In Appendix A, we list without proof some elementary properties of Λ-minimizers.

We now introduce the notion of cylindrical excess. The *height function* $\mathfrak{h} : \mathbb{H}^n \to \mathbb{R}$ is defined by $\mathfrak{h}(p) = p_1$, where p_1 is the first coordinate of $p = (p_1, \dots, p_{2n+1}) \in \mathbb{H}^n$. The set $\mathbb{W} = \{p \in \mathbb{H}^n : \mathfrak{h}(p) = 0\}$ is the vertical hyperplane passing through $0 \in \mathbb{H}^n$ and orthogonal to the left-invariant vector field X_1. The disk in $\mathbb{W}$ of radius $r > 0$ centred at $0 \in \mathbb{W}$ induced by the Korànyi norm is the set $D_r = \{p \in \mathbb{W} : \|p\|_K < r\}$. The intrinsic cylinder with central section D_r and height $2r$ is the set

$$C_r = D_r * (-r, r) \subset \mathbb{H}^n.$$

Here and in the sequel, we use the notation $D_r * (-r, r) = \{w * (se_1) \in \mathbb{H}^n : w \in D_r,\ s \in (-r, r)\big\}$, where $se_1 = (s, 0, \dots, 0) \in \mathbb{H}^n$. The cylinder C_r is comparable with the ball $B_r = \{\|p\|_K < r\}$. Namely, there exists a constant $k = k(n) \ge 1$ such that, for any $r > 0$, we have

$$B_{r/k} \subset C_r \subset B_{kr}. \tag{1-3}$$

By a rotation of the system of coordinates, it is enough to consider excess in cylinders with basis in $\mathbb{W}$ and axis X_1.

Definition 1.2 (cylindrical excess). Let $E \subset \mathbb{H}^n$ be a set with locally finite H-perimeter. The cylindrical excess of E at the point $0 \in \partial E$, at scale $r > 0$, and with respect to the direction $\nu = -X_1$ is defined as

$$\operatorname{Exc}(E, r, \nu) = \frac{1}{2r^{2n+1}} \int_{C_r} |\nu_E - \nu|_g^2 \, d\mu_E,$$

where μ_E is the H-perimeter measure of E and ν_E is its horizontal inner normal.

Theorem 1.3 (height estimate). *Let $n \geq 2$. There exist constants $\varepsilon_0 = \varepsilon_0(n) > 0$ and $c_0 = c_0(n) > 0$ with the following property: if $E \subset \mathbb{H}^n$ is a (Λ, r)-minimizer of H-perimeter in the cylinder C_{4k^2r}, $\Lambda r \leq 1$, $0 \in \partial E$, and*

$$\operatorname{Exc}(E, 4k^2 r, \nu) \leq \varepsilon_0,$$

then

$$\sup\{|\mathfrak{h}(p)| \in [0, \infty) : p \in \partial E \cap C_r\} \leq c_0\, r\, \operatorname{Exc}(E, 4k^2 r, \nu)^{1/(2(2n+1))}. \tag{1-4}$$

The constant $k = k(n)$ is the one in (1-3).

The estimate (1-4) does not hold when $n = 1$. In fact, there are sets $E \subset \mathbb{H}^1$ such that $\operatorname{Exc}(E, r, \nu) = 0$ but ∂E is not flat in $C_{\varepsilon r}$ for any $\varepsilon > 0$. See the conclusions of Proposition 3.7 in [Monti 2014]. Theorem 1.3 is proved in Section 3.

Besides local minimizers of H-perimeter, our interest in Λ-minimizers is also motivated by possible applications to isoperimetric sets. The height estimate is a first step in the regularity theory of Λ-minimizers of classical perimeter; we refer to [Maggi 2012, Part III] for a detailed account of the subject.

In order to state the slicing formula in its general form, we need the definition of a rectifiable set in $\mathbb{H}^n$ of codimension 1. We follow closely [Franchi et al. 2001], where this notion was first introduced.

The Riemannian and horizontal gradients of a function $f \in C^1(\mathbb{H}^n)$ are, respectively,

$$\nabla f = (X_1 f)X_1 + \cdots + (Y_n f)Y_n + (Tf)T,$$
$$\nabla_H f = (X_1 f)X_1 + \cdots + (Y_n f)Y_n.$$

We say that a continuous function $f \in C(\Omega)$, with $\Omega \subset \mathbb{H}^n$ an open set, is of class $C^1_H(\Omega)$ if the horizontal gradient $\nabla_H f$ exists in the sense of distributions and is represented by continuous functions $X_1 f, \ldots, Y_n f$ in Ω. A set $S \subset \mathbb{H}^n$ is an H-regular hypersurface if, for all $p \in S$, there exist $r > 0$ and a function $f \in C^1_H(B_r(p))$ such that $S \cap B_r(p) = \{q \in B_r(p) : f(q) = 0\}$ and $\nabla_H f(p) \neq 0$. Sets with H-regular boundary have locally finite H-perimeter.

For any $p = (z, t) \in \mathbb{H}^n$, let us define the box norm $\|p\|_\infty = \max\{|z|, |t|^{1/2}\}$ and the balls $U_r = \{q \in \mathbb{H}^n n \|q\|_\infty < r\}$ and $U_r(p) = p * U_r$ for $r > 0$. Let $E \subset \mathbb{H}^n$ be a set. For any $s \geq 0$ define the measure

$$\mathscr{S}^s(E) = \sup_{\delta > 0} \inf\left\{ c(n, s) \sum_{i \in \mathbb{N}} r_i^s : E \subset \bigcup_{i \in \mathbb{N}} U_{r_i}(p_i),\ r_i < \delta \right\}.$$

Above, $c(n, s) > 0$ is a normalization constant that we do not need to specify here. By Carathéodory's construction, $E \mapsto \mathscr{S}^s(E)$ is a Borel measure in $\mathbb{H}^n$. When $s = 2n + 2$, it turns out that $\mathscr{S}^{2n+2}$ is the Lebesgue measure $\mathscr{L}^{2n+1}$. Thus, the correct dimension to measure hypersurfaces is $s = 2n + 1$. In fact, if E is a set with locally finite H-perimeter in $\mathbb{H}^n$, then we have

$$\mu_E = \mathscr{S}^{2n+1} \llcorner \partial^* E, \tag{1-5}$$

where $\llcorner$ denotes restriction and $\partial^* E$ is the H-reduced boundary of E, namely the set of points $p \in \mathbb{H}^n$ such that $\mu_E(U_r(p)) > 0$ for all $r > 0$, $⨍_{U_r(p)} \nu_E \, d\mu_E \to \nu_E(p)$ as $r \to 0$, and $|\nu_E(p)|_g = 1$. The validity of formula (1-5) depends on the geometry of the balls $U_r(p)$; see [Magnani 2014]. We refer the reader to [Franchi et al. 2001] for more details on the H-reduced boundary.

Definition 1.4. A set $R \subset \mathbb{H}^n$ is $\mathscr{S}^{2n+1}$-rectifiable if there exists a sequence of H-regular hypersurfaces $(S_j)_{j\in\mathbb{N}}$ in $\mathbb{H}^n$ such that

$$\mathscr{S}^{2n+1}\Big(R \setminus \bigcup_{j\in\mathbb{N}} S_j\Big) = 0.$$

By the results of [Franchi et al. 2001], the H-reduced boundary $\partial^* E$ is $\mathscr{S}^{2n+1}$-rectifiable. Definition 1.4 is generalized in [Mattila et al. 2010], which studies the notion of an s-rectifiable set in $\mathbb{H}^n$ for any integer $1 \le s \le 2n+1$.

An H-regular surface S has a continuous horizontal normal ν_S that is locally defined up to the sign. This normal is given by the formula

$$\nu_S = \frac{\nabla_H f}{|\nabla_H f|_g}, \tag{1-6}$$

where f is a defining function for S. When $S = \partial E$ is the boundary of a smooth set, ν_S agrees with the horizontal normal ν_E. Then, for an $\mathscr{S}^{2n+1}$-rectifiable set $R \subset \mathbb{H}^n$, there is a unit horizontal normal $\nu_R : R \to H$ that is Borel regular. This normal is uniquely defined $\mathscr{S}^{2n+1}$-a.e. on R up to the sign; see Appendix B. However, (1-8) below does not depend on the sign.

In the following theorem, $\Omega \subset \mathbb{H}^n$ is an open set and $u \in C^\infty(\Omega)$ is a smooth function. For any $s \in \mathbb{R}$, we denote by $\Sigma^s = \{p \in \Omega : u(p) = s\}$ the level sets of u.

Theorem 1.5. *Let $R \subset \Omega$ be an $\mathscr{S}^{2n+1}$-rectifiable set. Then, for a.e. $s \in \mathbb{R}$ there exists a Radon measure μ_R^s on $R \cap \Sigma^s$ such that, for any Borel function $h : \Omega \to [0, \infty)$, the function*

$$s \mapsto \int_\Omega h \frac{|\nabla_H u|_g}{|\nabla u|_g}\, d\mu_R^s \tag{1-7}$$

is $\mathscr{L}^1$-measurable and we have the coarea formula

$$\int_{\mathbb{R}} \int_\Omega h \frac{|\nabla_H u|_g}{|\nabla u|_g}\, d\mu_R^s\, ds = \int_R h \sqrt{|\nabla_H u|_g^2 - \langle \nu_R, \nabla_H u\rangle_g^2}\, d\mathscr{S}^{2n+1}. \tag{1-8}$$

Theorem 1.5 is proved in Section 2. When $R \cap \Sigma^s$ is a regular subset of Σ^s, the measures μ_R^s are natural horizontal perimeters defined in Σ^s.

Coarea formulas in the Heisenberg group are known only for slicing of sets with positive Lebesgue measure; see [Magnani 2004; 2008]. Theorem 1.5 is, to our knowledge, the first example of slicing of lower-dimensional sets in a sub-Riemannian framework. Also, Theorem 1.5 is a nontrivial extension of the Riemannian coarea formula, because the set R and the slices $R \cap \Sigma^s$ need not be rectifiable in the standard sense; see [Kirchheim and Serra Cassano 2004]. We need the coarea formula (1-8) in the proof of Theorem 1.3; see Section 3C.

We conclude the introduction by stating a different but equivalent formulation of the coarea formula (1-8) that is closer to standard coarea formulas. This alternative formulation holds only when $n \ge 2$: when $n = 1$, the right-hand side in (1-9) might not be well defined; see Remark 2.11.

Theorem 1.6. *Let $\Omega \subset \mathbb{H}^n$, $n \geq 2$, be an open set, $u \in C^\infty(\Omega)$ be a smooth function, and $R \subset \Omega$ be an $\mathscr{S}^{2n+1}$-rectifiable set. Then, for any Borel function $h : \Omega \to [0, \infty)$,*

$$\int_{\mathbb{R}} \int_{\Omega} h \, d\mu_R^s \, ds = \int_R h |\nabla u|_g \sqrt{1 - \langle \nu_R, \nabla_H u / |\nabla_H u|_g \rangle_g^2} \, d\mathscr{S}^{2n+1}, \tag{1-9}$$

where μ_R^s are the measures given by Theorem 1.5.

2. Proof of the coarea formula

2A. ***Horizontal perimeter on submanifolds.*** Let $\Sigma \subset \mathbb{H}^n$ be a C^∞ hypersurface. We define the horizontal tangent bundle $H\Sigma$ by letting, for any $p \in \Sigma$,

$$H_p\Sigma = H_p \cap T_p\Sigma.$$

In general, the rank of $H\Sigma$ is not constant. This depends on the presence of *characteristic points* on Σ, i.e., points such that $H_p = T_p\Sigma$. For points $p \in \Sigma$ such that $H_p \neq T_p\Sigma$, we have $\dim(H_p\Sigma) = 2n - 1$.

We denote by σ_Σ the surface measure on Σ induced by the Riemannian metric g restricted to the tangent bundle $T\Sigma$.

Definition 2.1. Let $F \subset \Sigma$ be a Borel set and let $\Omega \subset \Sigma$ be an open set. We define the *H-perimeter of F in Ω*,

$$\mu_F^\Sigma(\Omega) = \sup\left\{ \int_F \operatorname{div}_g \varphi \, d\sigma_\Sigma : \varphi \in C_c^1(\Omega; H\Sigma), \ \|\varphi\|_g \leq 1 \right\}. \tag{2-10}$$

We say that the set $F \subset \Sigma$ has locally finite H-perimeter in Ω if $\mu_F^\Sigma(A) < \infty$ for any open set $A \Subset \Omega$.

By the Riesz theorem, if $F \subset \Sigma$ has locally finite H-perimeter in Ω, then the open sets mapping $A \mapsto \mu_F^\Sigma(A)$ extends to a Radon measure on Ω, called the *H-perimeter measure* of F.

Remark 2.2. If $F \subset \Sigma$ is an open set with smooth boundary, then, by the divergence theorem, we have, for any $\varphi \in C_c^1(\Omega; H\Sigma)$,

$$\int_F \operatorname{div}_g \varphi \, d\sigma_\Sigma = \int_{\partial F} \langle N_{\partial F}, \varphi \rangle_g \, d\lambda_{\partial F}, \tag{2-11}$$

where $N_{\partial F}$ is the Riemannian outer unit normal to ∂F and $d\lambda_{\partial F}$ is the Riemannian $(2n-1)$-dimensional volume form on ∂F induced by g.

From the sup definition (2-10) and from (2-11), we deduce that the H-perimeter measure of F has the representation

$$\mu_F^\Sigma = |N_{\partial F}^{H\Sigma}|_g \, \lambda_{\partial F},$$

where $N_{\partial F}^{H\Sigma} \in H\Sigma$ is the g-orthogonal projection of $N_{\partial F} \in T\Sigma$ onto $H\Sigma$.

This formula can be generalized as follows. We denote by $\mathscr{H}_g^{2n-1}$ the $(2n-1)$-dimensional Hausdorff measure in $\mathbb{H}^n$ induced by the metric g.

Lemma 2.3. *Let $F, \Omega \subset \Sigma$ be open sets and assume that there exists a compact set $N \subset \partial F$ such that $\mathscr{H}_g^{2n-1}(N) = 0$ and $(\partial F \setminus N) \cap \Omega$ is a smooth $(2n-1)$-dimensional surface. Then, we have*

$$\mu_F^\Sigma \llcorner \Omega = |N_{\partial F}^{H\Sigma}|_g \, \lambda_{\partial F \setminus N} \llcorner \Omega. \tag{2-12}$$

Proof. For any $\varepsilon > 0$ there exist points $p_i \in \mathbb{H}^n$ and radii $r_i \in (0, 1)$, $i = 1, \dots, M$, such that

$$N \subset \bigcup_{i=1}^{M} B_g(p_i, r_i) \quad \text{and} \quad \sum_{i=1}^{M} r_i^{2n-1} < \varepsilon,$$

where $B_g(p, r)$ denotes the ball in $\mathbb{H}^n$ with centre p and radius r with respect to the metric g. By a partition of unity argument, there exist functions f^ε, $g_i^\varepsilon \in C^\infty(\Omega; [0, 1])$, $i = 1, \dots, M$, such that:

(i) $f^\varepsilon + g_1^\varepsilon + \dots + g_M^\varepsilon = \chi_\Omega$;

(ii) $f^\varepsilon = 0$ on $\bigcup_{i=1}^M B_g(p_i, r_i/2)$;

(iii) for each i, the support of g_i^ε is contained in $B_g(p_i, r_i)$;

(iv) $|\nabla g_i^\varepsilon|_g \le C r_i^{-1}$ for a constant $C > 0$ independent of ε.

Hence, for any horizontal section $\varphi \in C_c^1(\Omega; H\Sigma)$, we have

$$\begin{aligned}\int_F \operatorname{div}_g \varphi \, d\sigma_\Sigma &= \int_F \operatorname{div}_g(f^\varepsilon \varphi) \, d\sigma_\Sigma + \sum_{i=1}^{M} \int_{F \cap B_g(p_i, r_i)} \operatorname{div}_g(g_i^\varepsilon \varphi) \, d\sigma_\Sigma \\ &= \int_{\partial F \setminus N} \langle f^\varepsilon \varphi, N_{\partial F} \rangle_g \, d\lambda_{\partial F \setminus N} + \sum_{i=1}^{M} \int_{F \cap B_g(p_i, r_i)} \operatorname{div}_g(g_i^\varepsilon \varphi) \, d\sigma_\Sigma,\end{aligned} \tag{2-13}$$

where, by (iv),

$$\left| \sum_{i=1}^{M} \int_{F \cap B_g(p_i, r_i)} \operatorname{div}_g(g_i^\varepsilon \varphi) \, d\sigma_\Sigma \right| \le \sum_{i=1}^{M} \int_{B_g(p_i, r_i)} (\| \operatorname{div}_g \varphi \|_{L^\infty} + C r_i^{-1}) \, d\sigma_\Sigma \le C' \sum_{i=1}^{M} r_i^{2n-1} \le C' \varepsilon \tag{2-14}$$

with a constant $C' > 0$ independent of ε.

Letting $\varepsilon \to 0$, we have $f^\varepsilon \to 1$ pointwise on $\partial F \setminus N$, by (i) and (iii). Then, from (2-13) and (2-14), we obtain

$$\int_F \operatorname{div}_g \varphi \, d\sigma_\Sigma = \int_{\partial F \setminus N} \langle \varphi, N_{\partial F} \rangle_g \, d\lambda_{\partial F \setminus N}$$

and claim (2-12) follows by standard arguments. □

2B. *Proof of Theorem 1.5.* Let $\Omega \subset \mathbb{H}^n$ be an open set and $u \in C^\infty(\Omega)$. By Sard's theorem, for a.e. $s \in \mathbb{R}$ the level set

$$\Sigma^s = \{p \in \Omega : u(p) = s\}$$

is a smooth hypersurface and, moreover, we have $\nabla u \neq 0$ on Σ^s.

Let $E \subset \mathbb{H}^n$ be a Borel set such that $E \cap \Sigma^s$ has (locally) finite H-perimeter in $\Omega \cap \Sigma^s$, in the sense of Definition 2.1. Then on $\Omega \cap \Sigma^s$ we have the H-perimeter measure $\mu^{\Sigma^s}_{E\cap\Sigma^s}$ induced by $E \cap \Sigma^s$. We shall use the notation

$$\mu^s_E = \mu^{\Sigma^s}_{E\cap\Sigma^s}$$

to denote a measure on Ω that is supported on $\Omega \cap \Sigma^s$.

We start with the following coarea formula in the smooth case, which is deduced from the Riemannian formula.

Lemma 2.4. *Let $\Omega \subset \mathbb{H}^n$ be an open set and $u \in C^\infty(\Omega)$. Let $E \subset \mathbb{H}^n$ be an open set with C^∞ boundary in Ω such that $\mu_E(\Omega) < \infty$. Then we have*

$$\int_{\mathbb{R}} \int_{\Omega} \frac{|\nabla_H u|_g}{|\nabla u|_g}\, d\mu^s_E\, ds = \int_{\Omega} \sqrt{|\nabla_H u|^2_g - \langle \nu_E, \nabla_H u\rangle^2_g}\, d\mu_E, \tag{2-15}$$

where μ_E is the H-perimeter measure of E and ν_E is its horizontal normal.

Proof. The integral in the left-hand side is well defined, because for a.e. $s \in \mathbb{R}$ there holds $\nabla u \neq 0$ on Σ^s. By the coarea formula for Riemannian manifolds — see, e.g., [Burago and Zalgaller 1988] — for any Borel function $h : \partial E \to [0, \infty]$ we have

$$\int_{\mathbb{R}} \int_{\partial E\cap\Sigma^s} h\, d\lambda_{\partial E\cap\Sigma^s}\, ds = \int_{\partial E} h |\nabla^{\partial E} u|_g\, d\sigma_{\partial E}, \tag{2-16}$$

where $\nabla^{\partial E} u$ is the tangential gradient of u on ∂E. Then we have

$$\nabla^{\partial E} u = \nabla u - \langle \nabla u, N_{\partial E}\rangle_g N_{\partial E} \quad \text{and} \quad |\nabla^{\partial E} u|_g = \sqrt{|\nabla u|^2_g - \langle \nabla u, N_{\partial E}\rangle^2_g}. \tag{2-17}$$

Step 1. Let us define the set

$$C = \left\{ p \in \partial E \cap \Omega : \nabla u(p) \neq 0 \text{ and } N_{\partial E}(p) = \pm \frac{\nabla u(p)}{|\nabla u(p)|_g} \right\}.$$

If $s \in \mathbb{R}$ is such that $\nabla u \neq 0$ on Σ^s, then $C \cap \Sigma^s$ is a closed set in Σ^s. Using the coarea formula (2-16) with the function $h = \chi_C$, we get

$$\int_{\mathbb{R}} \lambda_{\partial E\cap\Sigma^s}(C)\, ds = \int_C |\nabla^{\partial E} u|_g\, d\sigma_{\partial E} = 0,$$

because we have $\nabla^{\partial E} u = 0$ on C. In particular, we deduce that

$$C \cap \Sigma^s \text{ is a closed set in } \Sigma^s \quad \text{and} \quad \lambda_{\partial E\cap\Sigma^s}(C \cap \Sigma^s) = 0 \quad \text{for a.e. } s \in \mathbb{R}. \tag{2-18}$$

If $p \in \Sigma^s$ is a point such that $\nabla u(p) \neq 0$ and $p \notin C$, then Σ^s is a smooth hypersurface in a neighbourhood of p and $E^s = E \cap \Sigma^s$ is a domain in Σ^s with smooth boundary in a neighbourhood of p. Moreover, we have $(\partial E \cap \Sigma^s) \setminus C = \partial E^s \setminus C$. Then, from (2-18) and Lemma 2.3 we conclude that for a.e. $s \in \mathbb{R}$ we have

$$\mu^s_E = |N^{H\Sigma^s}_{\partial E^s}|_g \lambda_{\partial E^s}. \tag{2-19}$$

By (2-18) and (2-19),

$$\mu_E^s(C\cap\Sigma^s)=\int_{C\cap\Sigma^s}|N_{\partial E^s}^{H\Sigma^s}|_g\,d\lambda_{\partial E^s}=0\quad\text{for a.e. } s\in\mathbb{R}. \tag{2-20}$$

Step 2. We prove (2-15) by plugging into (2-16) the Borel function $h:\partial E\to[0,\infty]$,

$$h=\begin{cases}\dfrac{|N_{\partial E}^H|_g\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}}{|\nabla u|_g\sqrt{1-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g^2}} & \text{on } \partial E\setminus(C\cup\{\nabla u=0\}),\\ 0 & \text{on } C\cup\{\nabla u=0\}.\end{cases}$$

Above, $N_{\partial E}^H$ is the projection of the Riemannian normal $N_{\partial E}$ onto H and ν_E is the horizontal normal. Namely, we have

$$N_{\partial E}^H=N_{\partial E}-\langle N_{\partial E},T\rangle_g T\quad\text{and}\quad \nu_E=\frac{N_{\partial E}^H}{|N_{\partial E}^H|_g}.$$

The H-perimeter measure of E is

$$\mu_E=|N_{\partial E}^H|_g\sigma_{\partial E}. \tag{2-21}$$

Using (2-17) and (2-21), we find

$$\begin{aligned}\int_{\partial E}h|\nabla^{\partial E}u|\,d\sigma_{\partial E}&=\int_{\partial E\setminus(C\cup\{\nabla u=0\})}|N_{\partial E}^H|_g\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\sigma_{\partial E}\\&=\int_{\partial E\setminus(C\cup\{\nabla u=0\})}\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E\\&=\int_{\partial E}\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E,\end{aligned} \tag{2-22}$$

where the last equality is justified by the fact that if $p\in C\cup\{\nabla u=0\}$ then

$$\sqrt{|\nabla_H u(p)|_g^2-\langle\nu_E(p),\nabla_H u(p)\rangle_g^2}=0.$$

For a.e. $s\in\mathbb{R}$, we have $\nabla u\neq 0$ on Σ^s. Using (2-21) and the fact that $h=0$ on $C\cup\{\nabla_H u=0\}$, letting $\Lambda^s=(\partial E\cap\Sigma^s)\setminus(C\cup\{\nabla_H u=0\})$, we obtain

$$\begin{aligned}\int_{\mathbb{R}}\int_{\partial E\cap\Sigma^s}h\,d\lambda_{\partial E^s}\,ds&=\int_{\mathbb{R}}\int_{\Lambda^s}\frac{|N_{\partial E}^H|_g\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}}{|\nabla u|_g\sqrt{1-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g^2}}\,d\lambda_{\partial E^s}\,ds\\&=\int_{\mathbb{R}}\int_{\Lambda^s}\frac{|\nabla_H u|_g}{|\nabla u|_g}\vartheta^s\,d\lambda_{\partial E^s}\,ds,\end{aligned} \tag{2-23}$$

where we let

$$\vartheta^s=\frac{\sqrt{|N_{\partial E}^H|_g^2-\langle N_{\partial E}^H,\nabla_H u/|\nabla_H u|_g\rangle_g^2}}{\sqrt{1-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g^2}}.$$

We will prove in Step 3 that, for any $s\in\mathbb{R}$ such that $\nabla u\neq 0$ on Σ^s,

$$\vartheta^s=|N_{\partial E^s}^{H\Sigma^s}|_g\quad\text{on } \Lambda^s. \tag{2-24}$$

Using (2-24), (2-19), and (2-20), formula (2-23) becomes

$$\begin{aligned}\int_{\mathbb{R}}\int_{\partial E\cap\Sigma^s} h\,d\lambda_{\partial E\cap\Sigma^s}\,ds &= \int_{\mathbb{R}}\int_{\Lambda^s}\frac{|\nabla_H u|_g}{|\nabla u|_g}|N_{\partial E^s}^{H\Sigma^s}|_g\,d\lambda_{\partial E^s}\,ds\\ &= \int_{\mathbb{R}}\int_{\Lambda^s}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds\\ &= \int_{\mathbb{R}}\int_{\partial E\cap\Sigma^s}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds.\end{aligned}\tag{2-25}$$

The proof is complete, because (2-15) follows from (2-16), (2-22), and (2-25).

Step 3. We prove claim (2-24). Let us introduce the vector field W in $\Omega\setminus\{\nabla_H u=0\}$,

$$W=\frac{Tu}{|\nabla u|_g}\frac{\nabla_H u}{|\nabla_H u|_g}-\frac{|\nabla_H u|_g}{|\nabla u|_g}T.$$

It can be checked that $|W|_g=1$ and $Wu=0$. In particular, for a.e. s we have $W\in T\Sigma^s$. Moreover, W is g-orthogonal to $H\Sigma^s$ because any vector in $H\Sigma^s$ is orthogonal both to $\nabla_H u$ and to T. It follows that

$$N_{\partial E^s}^{H\Sigma^s}=N_{\partial E^s}-\langle N_{\partial E^s},W\rangle_g$$

and, in particular,

$$|N_{\partial E^s}^{H\Sigma^s}|_g^2=1-\langle N_{\partial E^s},W\rangle_g^2.$$

Starting from the formula

$$N_{\partial E^s}=\frac{N_{\partial E}-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g\nabla u/|\nabla u|_g}{|N_{\partial E}-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g\nabla u/|\nabla u|_g|_g}=\frac{N_{\partial E}-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g\nabla u/|\nabla u|_g}{\sqrt{1-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g^2}},$$

we find

$$|N_{\partial E^s}^{H\Sigma^s}|_g^2=\frac{M}{1-\langle N_{\partial E},\nabla u/|\nabla u|_g\rangle_g^2},$$

where we let

$$M=1-\left\langle N_{\partial E},\frac{\nabla u}{|\nabla u|_g}\right\rangle_g^2-\left\langle N_{\partial E}-\left\langle N_{\partial E},\frac{\nabla u}{|\nabla u|_g}\right\rangle_g\frac{\nabla u}{|\nabla u|_g},W\right\rangle_g^2.$$

We claim that, on the open set $\{\nabla_H u\neq 0\}$,

$$M=|N_{\partial E}^H|_g^2-\left\langle N_{\partial E}^H,\frac{\nabla_H u}{|\nabla_H u|_g}\right\rangle_g^2,\tag{2-26}$$

and formula (2-24) follows from (2-26). Using the identity $\nabla u=\nabla_H u+(Tu)T$ and the orthogonality

$$\left\langle N_{\partial E}-\left\langle N_{\partial E},\frac{\nabla u}{|\nabla u|_g}\right\rangle_g\frac{\nabla u}{|\nabla u|_g},\nabla u\right\rangle_g=0,$$

we find

$$
\begin{aligned}
M &= 1 - \left\langle N_{\partial E}, \frac{\nabla_H u + (Tu)T}{|\nabla u|_g} \right\rangle_g^2 - \left(\frac{Tu}{|\nabla u|_g} \left\langle N_{\partial E}, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g - \frac{|\nabla_H u|_g}{|\nabla u|_g} \langle N_{\partial E}, T \rangle_g \right)^2 \\
&= 1 - \left\langle N_{\partial E}, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g^2 \frac{|\nabla_H u|_g^2 + (Tu)^2}{|\nabla u|_g^2} - \langle N_{\partial E}, T \rangle_g^2 \frac{|\nabla_H u|_g^2 + (Tu)^2}{|\nabla u|_g^2} \\
&= 1 - \left\langle N_{\partial E}, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g^2 - \langle N_{\partial E}, T \rangle_g^2 \\
&= 1 - \langle N_{\partial E}, T \rangle_g^2 - \left(\left\langle N_{\partial E}, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g - \left\langle \langle N_{\partial E}, T \rangle_g T, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g \right)^2 \\
&= |N_{\partial E}^H|_g^2 - \left\langle N_{\partial E}^H, \frac{\nabla_H u}{|\nabla_H u|_g} \right\rangle_g^2 .
\end{aligned} \tag{2-27}
$$

This ends the proof. □

We prove a coarea inequality:

Proposition 2.5. *Let $\Omega \subset \mathbb{H}^n$ be an open set, $u \in C^\infty(\Omega)$ a smooth function, $E \subset \mathbb{H}^n$ a set with finite H-perimeter in Ω, and let $h : \partial E \to [0, \infty]$ be a Borel function. Then we have*

$$
\int_{\mathbb{R}} \int_{\Omega} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \le \int_{\Omega} h \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E . \tag{2-28}
$$

Proof. The coarea inequality (2-28) follows from the smooth case of Lemma 2.4 by an approximation and lower semicontinuity argument.

Step 1. By [Franchi et al. 1996, Theorem 2.2.2], there exists a sequence of smooth sets $(E_j)_{j \in \mathbb{N}}$ in Ω such that

$$
\chi_{E_j} \xrightarrow{L^1(\Omega)} \chi_E \quad \text{as } j \to \infty \qquad \text{and} \qquad \lim_{j \to \infty} \mu_{E_j}(\Omega) = \mu(\Omega).
$$

By a straightforward adaptation of the proof of [Ambrosio et al. 2000, Proposition 3.13], we also have that $\nu_{E_j} \mu_{E_j} \to \nu_E \mu_E$ weakly* in Ω. Namely, for any $\psi \in C_c(\Omega; H)$,

$$
\lim_{j \to \infty} \int_{\Omega} \langle \psi, \nu_{E_j} \rangle_g \, d\mu_{E_j} = \int_{\Omega} \langle \psi, \nu_E \rangle_g \, d\mu_E .
$$

Let $A \Subset \Omega$ be an open set such that $\lim_{j \to \infty} \mu_{E_j}(A) = \mu_E(A)$. By Reshetnyak's continuity theorem (see, e.g., [Ambrosio et al. 2000, Theorem 2.39]), we have

$$
\lim_{j \to \infty} \int_A f(p, \nu_{E_j}(p)) \, d\mu_{E_j} = \int_A f(p, \nu_E(p)) \, d\mu_E
$$

for any continuous and bounded function f. In particular,

$$
\lim_{j \to \infty} \int_A \sqrt{|\nabla_H u|_g^2 - \langle \nu_{E_j}, \nabla_H u \rangle_g^2} \, d\mu_{E_j} = \int_A \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E . \tag{2-29}
$$

Step 2. Let $(E_j)_{j\in\mathbb{N}}$ be the sequence introduced in Step 1. Then, for a.e. $s\in\mathbb{R}$, we have

$$\nabla u\neq 0 \quad \text{on } \Sigma^s \qquad \text{and} \qquad \chi_{E_j}\to\chi_E \quad \text{in } L^1(\Sigma^s,\sigma_{\Sigma^s}) \text{ as } j\to\infty.$$

In particular, for any such s and for any open set $A\subset\Sigma^s\cap\Omega$,

$$\mu_E^s(A)\le\liminf_{j\to\infty}\mu_{E_j}^s(A).$$

From Fatou's lemma and the continuity of $|\nabla_H u|_g/|\nabla u|_g$ on Σ^s, it follows that

$$\begin{aligned}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s&=\int_0^\infty\mu_E^s\Bigl(\Bigl\{p\in A:\frac{|\nabla_H u|_g}{|\nabla u|_g}(p)>t\Bigr\}\Bigr)\,dt\\&\le\int_0^\infty\liminf_{j\to\infty}\mu_{E_j}^s\Bigl(\Bigl\{p\in A:\frac{|\nabla_H u|_g}{|\nabla u|_g}(p)>t\Bigr\}\Bigr)\,dt\\&\le\liminf_{j\to\infty}\int_0^\infty\mu_{E_j}^s\Bigl(\Bigl\{p\in A:\frac{|\nabla_H u|_g}{|\nabla u|_g}(p)>t\Bigr\}\Bigr)\,dt\\&=\liminf_{j\to\infty}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_{E_j}^s.\end{aligned}$$

Using again Fatou's lemma and Lemma 2.4,

$$\begin{aligned}\int_{\mathbb{R}}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds&\le\int_{\mathbb{R}}\liminf_{j\to\infty}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_{E_j}^s\,ds\\&\le\liminf_{j\to\infty}\int_{\mathbb{R}}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_{E_j}^s\,ds\\&=\liminf_{j\to\infty}\int_A\sqrt{|\nabla_H u|_g^2-\langle\nu_{E_j},\nabla_H u\rangle_g^2}\,d\mu_{E_j}.\end{aligned}$$

This, together with (2-29), gives

$$\int_{\mathbb{R}}\int_A\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds\le\int_A\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E.$$

Step 3. Any open set $A\subset\Omega$ can be approximated by a sequence $(A_k)_{k\in\mathbb{N}}$ of open sets such that

$$A_k\Subset\Omega,\quad A_k\subset A_{k+1},\quad\bigcup_{k=1}^\infty A_k=A\quad\text{and}\quad\mu_E(\partial A_k)=0.$$

In particular, for each $k\in\mathbb{N}$, we have

$$\liminf_{j\to\infty}\mu_{E_j}(A_k)\le\limsup_{j\to\infty}\mu_{E_j}(\bar A_k)\le\mu_E(\bar A_k)=\mu_E(A_k)\le\liminf_{j\to\infty}\mu_{E_j}(A_k).$$

Hence, the inequalities are equalities, i.e., $\mu_E(A_k)=\lim_{j\to\infty}\mu_{E_j}(A_k)$. By Step 2, for any $k\in\mathbb{N}$,

$$\int_{\mathbb{R}}\int_{A_k}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds\le\int_{A_k}\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E.$$

By monotone convergence, letting $k \to \infty$ we obtain, for any open set $A \subset \Omega$,

$$\int_{\mathbb{R}} \int_A \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \le \int_A \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E.$$

By a standard approximation argument, it is enough to prove (2-28) for the characteristic function $h = \chi_B$ of a Borel set $B \subset \partial E$. Since the measure $\sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, \mu_E$ is a Radon measure on ∂E, there exists a sequence of open sets B_j such that $B \subset B_j$ for each $j \in \mathbb{N}$ and

$$\lim_{j \to \infty} \int_{B_j} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E = \int_B \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E.$$

Therefore, we have

$$\begin{aligned} \int_{\mathbb{R}} \int_B \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds &\le \liminf_{j \to \infty} \int_{\mathbb{R}} \int_{B_j} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \\ &\le \lim_{j \to \infty} \int_{B_j} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E \\ &= \int_B \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E, \end{aligned}$$

and this concludes the proof. □

In the next step, we prove an approximate coarea formula for sets E such that the boundary ∂E is an H-regular surface.

Lemma 2.6. *Let $\Omega \subset \mathbb{H}^n$ be an open set, $u \in C^\infty(\Omega)$ a smooth function, $E \subset \mathbb{H}^n$ an open set such that $\partial E \cap \Omega$ is an H-regular hypersurface, and $\bar{p} \in \partial E \cap \Omega$ a point such that*

$$\nabla_H u(\bar{p}) \neq 0 \quad \textit{and} \quad \nu_E(\bar{p}) \neq \pm \frac{\nabla_H u(\bar{p})}{|\nabla_H u(\bar{p})|_g}.$$

Then, for any $\varepsilon > 0$, there exists $\bar{r} = \bar{r}(\bar{p}, \varepsilon) > 0$ such that $B_{\bar{r}}(\bar{p}) \subset \Omega$ and, for any $r \in (0, \bar{r})$,

$$\begin{aligned} (1 - \varepsilon) \int_{B_r(\bar{p})} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E &\le \int_{\mathbb{R}} \int_{B_r(\bar{p})} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \\ &\le (1 + \varepsilon) \int_{B_r(\bar{p})} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E. \end{aligned}$$

Proof. We can, without loss of generality, assume that $\bar{p} = 0$ and $u(0) = 0$. We divide the proof into several steps.

Step 1: preliminary considerations. The horizontal vector field $V_{2n} = \nabla_H u / |\nabla_H u|_g$ is well defined in a neighbourhood $\Omega_\varepsilon \subset \mathbb{H}^n$ of 0. For any $s \in \mathbb{R}$, the hypersurface $\Sigma^s = \{p \in \Omega : u(p) = s\}$ is smooth in Ω_ε because $\nabla_H u \neq 0$ on Ω_ε.

There are horizontal vector fields $V_1, \dots, V_{2n-1}$ on Ω_ε such that $V_1, \dots, V_{2n}$ is a g-orthonormal frame. In particular, we have $V_j u = 0$ for $j = 1, \dots, 2n - 1$, i.e.,

$$H_p \Sigma^s = \operatorname{span}\{V_1(p), \dots, V_{2n-1}(p)\} \quad \text{for all } p \in \Sigma^s \cap \Omega_\varepsilon. \tag{2-30}$$

Possibly shrinking Ω_ε, reordering $\{V_j\}_{j=1,\dots,2n-1}$, and changing the sign of V_1, we can assume (see [Vittone 2012, Lemmas 4.3 and 4.4]) that there exist a function $f:\Omega_\varepsilon\to\mathbb{R}$ and a number $\delta>0$ such that:

(a) $f\in C^1_H(\Omega_\varepsilon)\cap C^\infty(\Omega_\varepsilon\setminus\partial E)$;

(b) $E\cap\Omega_\varepsilon=\{p\in\Omega_\varepsilon : f(p)>0\}$;

(c) $V_1 f\ge\delta>0$ on Ω_ε.

By [Vittone 2012, Remark 4.7], we also have $\nu_E=\nabla_H f/|\nabla_H f|_g$ on $\partial E\cap\Omega_\varepsilon$.

Step 2: change of coordinates. Let $S\subset\mathbb{H}^n$ be a $(2n-1)$-dimensional smooth submanifold such that:

(i) $0\in S$.

(ii) $S\subset\Sigma^0\cap\Omega_\varepsilon$. In particular, ∇u is g-orthogonal to S.

(iii) $V_1(0)$ is g-orthogonal to S at 0.

(iv) There exists a diffeomorphism $H:U\to\mathbb{H}^n$, where $U\subset\mathbb{R}^{2n-1}$ is an open set with $0\in U$, such that $H(0)=0$ and $H(U)=S\cap\Omega_\varepsilon$.

(v) The area element JH of H satisfies $JH(0)=1$. Namely,

$$JH(0)=\lim_{r\to 0}\frac{\lambda_S(H(B^E_r))}{\mathscr{L}^{2n-1}(B^E_r)}=1,$$

where $B^E_r=\{p\in\mathbb{R}^{2n-1} : |p|<r\}$ is a Euclidean ball and λ_S is the Riemannian $(2n-1)$-volume measure on S induced by g.

For small enough $a,b>0$, and possibly shrinking U and Ω_ε, the mapping $G:(-a,a)\times(-b,b)\times U\to\mathbb{H}^n$,

$$G(v,z,w)=\exp(vV_1)\exp\Bigl(z\frac{\nabla u}{|\nabla u|_g^2}\Bigr)(H(w))$$

is a diffeomorphism from $\widetilde{\Omega}_\varepsilon=(-a,a)\times(-b,b)\times U$ onto Ω_ε. The differential of G satisfies

$$dG\Bigl(\frac{\partial}{\partial v}\Bigr)=V_1\quad\text{and}\quad dG(0)\Bigl(\frac{\partial}{\partial z}\Bigr)=\frac{\nabla u(0)}{|\nabla u(0)|_g^2}.$$

Moreover, the tangent space $T_0S=\operatorname{Im} dH(0)$ is g-orthogonal to $V_1(0)$ and $\nabla u(0)/|\nabla u(0)|_g^2$. We denote by G_z the restriction of G to $(-a,a)\times\{z\}\times U$, i.e., $G_z(v,w)=G(v,z,w)$. From the above considerations, we deduce that the area elements of G and G_0 satisfy

$$JG(0)=\frac{1}{|\nabla u(0)|_g}\quad\text{and}\quad JG_0(0)=1.$$

Then, possibly shrinking $\widetilde{\Omega}_\varepsilon$ further, we have

$$(1-\varepsilon)JG(v,z,w)\le\frac{JG_z(v,w)}{|\nabla u\circ G(v,z,w)|_g}\le(1+\varepsilon)JG(v,z,w)\tag{2-31}$$

for all $(v,z,w)\in\widetilde{\Omega}_\varepsilon$.

For $j = 1, \dots, 2n$, we define on $\widetilde{\Omega}_\varepsilon$ the vector fields $\widetilde{V}_j = (dG)^{-1}(V_j)$. By the definition of G, we have $\widetilde{V}_1 = \partial/\partial v$. We also define $\tilde{u} = u \circ G \in C^\infty(\widetilde{\Omega}_\varepsilon)$, $\tilde{f} = f \circ G : \widetilde{\Omega}_\varepsilon \to \mathbb{R}$, and $\widetilde{E} = G^{-1}(E)$. Then:

(1) $\widetilde{E} = \{q \in \widetilde{\Omega}_\varepsilon : \tilde{f}(q) > 0\}$.

(2) $\tilde{f} \in C^\infty(\widetilde{\Omega}_\varepsilon \setminus \partial \widetilde{E})$.

(3) The derivative $\widetilde{V}_j \tilde{f}$ is defined in the sense of distributions with respect to the measure $\mu = JG\mathscr{L}^{2n+1}$. Namely, for all $\psi \in C_c^\infty(\widetilde{\Omega}_\varepsilon)$, we have

$$\int_{\widetilde{\Omega}_\varepsilon} (\widetilde{V}_j \tilde{f})\, \psi \, d\mu = -\int_{\widetilde{\Omega}_\varepsilon} \tilde{f}\, \widetilde{V}_j^* \psi \, d\mu,$$

where $\widetilde{V}_j^*$ is the adjoint operator of $\widetilde{V}_j$ with respect to μ. Then $\widetilde{V}_j \tilde{f} = (V_j f) \circ G$ and so $\widetilde{V}_j \tilde{f}$ is a continuous function for any $j = 1, \dots, 2n$. In particular, $\widetilde{V}_1 \tilde{f} = \partial_v \tilde{f} \geq \delta > 0$.

Step 3: approximate coarea formula. We follow the argument of [Vittone 2012, Propositions 4.1 and 4.5]; see also Remark 4.7 therein.

Possibly shrinking $\widetilde{\Omega}_\varepsilon$ and Ω_ε, there exists a continuous function $\phi : (-b, b) \times U \to (-a, a)$ such that:

(A) $\partial \widetilde{E} \cap \widetilde{\Omega}_\varepsilon$ is the graph of ϕ. Namely, letting $\Phi : (-b, b) \times U \to \mathbb{R}^{2n+1}$, $\Phi(z, w) = (\phi(z, w), z, w)$, we have

$$\partial \widetilde{E} \cap \widetilde{\Omega}_\varepsilon = \Phi((-b, b) \times U).$$

(B) The measure μ_E is

$$\mu_E \llcorner \Omega_\varepsilon = (G \circ \Phi)_\# \left(\left(\frac{|\widetilde{V} \tilde{f}|}{\widetilde{V}_1 \tilde{f}} JG \right) \circ \Phi \; \mathscr{L}^{2n} \llcorner ((-b, b) \times U) \right), \tag{2-32}$$

where $(G \circ \Phi)_\#$ denotes the push-forward and

$$|\widetilde{V} \tilde{f}| = \left(\sum_{j=1}^{2n} (\widetilde{V}_j \tilde{f})^2 \right)^{\frac{1}{2}}.$$

Using $V_1 u = 0$ and $u \circ H = 0$ (this follows from $H(U) = S \cap \Omega_\varepsilon \subset \Sigma^0 \cap \Omega_\varepsilon$), we obtain

$$\begin{aligned}\tilde{u}(v, z, w) = u(G(v, z, w)) = u\left(\exp(v V_1) \exp\left(z \frac{\nabla u}{|\nabla u|_g^2} \right)(H(w)) \right) &= u\left(\exp\left(z \frac{\nabla u}{|\nabla u|_g^2} \right)(H(w)) \right) \\ &= z + u(H(w)) = z.\end{aligned}$$

In particular, from $\tilde{u} = u \circ G$, we deduce that

$$G^{-1}(\Sigma^s \cap \Omega_\varepsilon) = (-a, a) \times \{s\} \times U.$$

We denote by JG_s the Jacobian (area element) of G_s. We also define the restriction $\Phi_s : U \to \mathbb{R}^{2n+1}$, $\Phi_s(w) = \Phi(s, w)$, for any $s \in (-b, b)$.

By (2-30), for any $s \in \mathbb{R}$, the measure $\mu_E^s = \mu_{E \cap \Sigma_s}^{\Sigma^s}$ is the horizontal perimeter of $E \cap \Sigma^s$ with respect to the Carnot–Carathéodory structure induced by the family $V_1, \dots, V_{2n-1}$ on Σ^s. We can repeat the argument that led to (2-32) to obtain

$$\mu_E^s \llcorner \Omega_\varepsilon = (G\circ\Phi_s)_\#\left(\left(\frac{|\widetilde{V}'\tilde f|}{\widetilde{V}_1\tilde f}\,JG_s\right)\circ\Phi_s\mathscr{L}^{2n-1}\llcorner U\right), \tag{2-33}$$

where $\widetilde{V}'\tilde f=(\widetilde{V}_1\tilde f,\dots,\widetilde{V}_{2n-1}\tilde f)$. We omit the details of the proof of (2-33). The proof is a line-by-line repetition of Proposition 4.5 in [Vittone 2012] with the sole difference that now the horizontal perimeter is defined in a curved manifold.

Let us fix $\bar r>0$ such that $B_{\bar r}\subset\Omega_\varepsilon$ and, for any $r\in(0,\bar r)$, let

$$A_{s,r}=\{w\in U: G(0,s,w)\in B_r\}\quad\text{and}\quad A_r=\{(s,w)\in(-b,b)\times U: w\in A_{s,r}\}.$$

By the Fubini–Tonelli theorem and (2-33), the function

$$s\mapsto\int_{B_r}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s=\int_{A_{s,r}}\left(\frac{|\nabla_H u|_g}{|\nabla u|_g}\circ G\right)\left(\frac{|\widetilde{V}'\tilde f|}{\widetilde{V}_1\tilde f}\,JG_s\right)\circ\Phi_s\,d\mathscr{L}^{2n-1} \tag{2-34}$$

is $\mathscr{L}^1$-measurable. Here and hereafter, the composition $\cdot\circ\Phi_s$ acts on the product. Thus, from the Fubini–Tonelli theorem and (2-31), we obtain

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{B_r}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds\\
&\qquad=\int_{\mathbb{R}}\int_{A_{s,r}}\left(\frac{|\nabla_H u|_g}{|\nabla u|_g}\circ G\right)\left(\frac{|\widetilde{V}'\tilde f|}{\widetilde{V}_1\tilde f}\,JG_s\right)\circ\Phi_s(w)\,d\mathscr{L}^{2n-1}(w)\,ds\\
&\qquad=\int_{A_r}(|\nabla_H u|_g\circ G)\left(\frac{|\widetilde{V}'\tilde f|}{\widetilde{V}_1\tilde f}\,\frac{JG_s}{|\nabla u|_g\circ G}\right)\circ\Phi(s,w)\,d\mathscr{L}^{2n}(s,w)\\
&\qquad\le(1+\varepsilon)\int_{A_r}(|\nabla_H u|_g\circ G)\left(\frac{|\widetilde{V}\tilde f|}{\widetilde{V}_1\tilde f}\sqrt{1-(\widetilde{V}_{2n}\tilde f)^2/|\widetilde{V}\tilde f|^2}\,JG\right)\circ\Phi(s,w)\,d\mathscr{L}^{2n}(s,w).
\end{aligned} \tag{2-35}$$

From the identity

$$\frac{\widetilde{V}_{2n}\tilde f}{|\widetilde{V}\tilde f|}=\frac{V_{2n}f}{|\nabla_H f|_g}\circ G=\left\langle\frac{\nabla_H u}{|\nabla_H u|_g},\frac{\nabla_H f}{|\nabla_H f|_g}\right\rangle_g\circ G=\left\langle\frac{\nabla_H u}{|\nabla_H u|_g},\nu_E\right\rangle_g\circ G \tag{2-36}$$

and from (2-32), we deduce that

$$\begin{aligned}
\int_{\mathbb{R}}\int_{B_r}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds&\le(1+\varepsilon)\int_{B_r}|\nabla_H u|_g\sqrt{1-\langle\nabla_H u/|\nabla_H u|_g,\nu_E\rangle_g^2}\,d\mu_E\\
&=(1+\varepsilon)\int_{B_r}\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E.
\end{aligned} \tag{2-37}$$

In a similar way, we obtain

$$\int_{\mathbb{R}}\int_{B_r}\frac{|\nabla_H u|_g}{|\nabla u|_g}\,d\mu_E^s\,ds\ge(1-\varepsilon)\int_{B_r}\sqrt{|\nabla_H u|_g^2-\langle\nu_E,\nabla_H u\rangle_g^2}\,d\mu_E.$$

This concludes the proof. ☐

We can now prove the coarea formula for H-regular boundaries.

Proposition 2.7. *Let $\Omega \subset \mathbb{H}^n$ be an open set, $u \in C^\infty(\Omega)$, and let $E \subset \mathbb{H}^n$ be an open domain such that $\partial E \cap \Omega$ is an H-regular hypersurface. Then*

$$\int_{\mathbb{R}} \int_{\Omega} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds = \int_{\Omega} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E. \tag{2-38}$$

Proof. Let us define the set

$$A = \left\{ p \in \partial E \cap \Omega : \nabla_H u(p) \neq 0 \text{ and } \nu_E(p) \neq \pm \frac{\nabla_H u(p)}{|\nabla_H u(p)|_g} \right\}.$$

The set A is relatively open in $\partial E \cap \Omega$. Let $\varepsilon > 0$ be fixed. Since the measure μ_E is locally doubling on $\partial E \cap \Omega$ (see, e.g., [Vittone 2012, Corollary 4.13]), by Lemma 2.6 and the Vitali covering theorem (see, e.g., [Heinonen 2001, Theorem 1.6]) there exists a countable (or finite) collection of balls $B_{r_i}(p_i)$, $i \in \mathbb{N}$, such that:

(i) for any $i \in \mathbb{N}$ we have $p_i \in A$ and $0 < r_i < \bar{r}(p_i, \varepsilon)$, where $\bar{r}$ is as in the statement of Lemma 2.6;

(ii) the balls $B_{r_i}(p_i)$ are contained in A and pairwise disjoint;

(iii) $\mu_E\big(A \setminus \bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)\big) = 0$.

It follows that we have

$$\begin{aligned} \int_{\mathbb{R}} \int_{\bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds &\le (1+\varepsilon) \int_{\bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E \\ &= (1+\varepsilon) \int_A \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E \\ &= (1+\varepsilon) \int_\Omega \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E. \end{aligned} \tag{2-39}$$

The last equality follows from the fact that $\sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} = 0$ outside A. In the same way, one also obtains

$$\int_{\mathbb{R}} \int_{\bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \ge (1-\varepsilon) \int_\Omega \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E. \tag{2-40}$$

Moreover, by Proposition 2.5,

$$\int_{\mathbb{R}} \int_{\Omega \setminus \bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)} \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \le \int_{\Omega \setminus \bigcup_{i \in \mathbb{N}} B_{r_i}(p_i)} \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E = 0.$$

In particular, the integral on the left-hand side of the last inequality is 0 and, by (2-39) and (2-40), we obtain

$$(1-\varepsilon) \int_\Omega \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E \le \int_{\mathbb{R}} \int_\Omega \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds \le (1+\varepsilon) \int_\Omega \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E.$$

Since $\varepsilon > 0$ is arbitrary, this concludes the proof. $\square$

By a standard approximation argument, we also have this extension of the coarea formula (2-38):

Proposition 2.8. *Let $\Omega \subset \mathbb{H}^n$ be an open set, $u \in C^\infty(\Omega)$, and let E be an open domain such that $\partial E \cap \Omega$ is an H-regular hypersurface. Then, for any Borel function $h : \partial E \to [0,\infty)$,*

$$\int_{\mathbb{R}} \int_{\Omega} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_E^s \, ds = \int_{\Omega} h \sqrt{|\nabla_H u|_g^2 - \langle \nu_E, \nabla_H u \rangle_g^2} \, d\mu_E.$$

Our next step is to prove the coarea formula for $\mathscr{S}^{2n+1}$-rectifiable sets.

Lemma 2.9. *Let $R \subset \mathbb{H}^n$ be an $\mathscr{S}^{2n+1}$-rectifiable set. Then, there exists a Borel $\mathscr{S}^{2n+1}$-rectifiable set $R' \subset \mathbb{H}^n$ such that $\mathscr{S}^{2n+1}(R \Delta R') = 0$.*

Proof. By assumption, there exist a $\mathscr{S}^{2n+1}$-negligible set N and H-regular hypersurfaces $S_j \subset \mathbb{H}^n$, $j \in \mathbb{N}$, such that

$$R \subset N \cup \bigcup_{j=1}^{\infty} S_j.$$

It is proved in [Franchi et al. 2001; Ambrosio et al. 2006] that (up to a localization argument), for any $j \in \mathbb{N}$, there exist an open set $U_j \subset \mathbb{R}^{2n}$, a homeomorphism $\Phi_j : U_j \to S_j$, and a continuous function $\rho_j : U_j \to [1,\infty)$ such that $\mathscr{S}^{2n+1} \llcorner S_j = \Phi_{j\#}(\rho_j \mathscr{L}^{2n} \llcorner U_j)$. Since the Lebesgue measure $\mathscr{L}^{2n}$ is a complete Borel measure, for any $j \in \mathbb{N}$ there exists a Borel set $T_j \subset U_j$ such that

$$\mathscr{L}^{2n}(T_j \Delta \Phi_j^{-1}(R \cap S_j)) = 0.$$

In particular, the Borel set

$$R' = \bigcup_{j=1}^{\infty} \Phi_j(T_j)$$

is $\mathscr{S}^{2n+1}$-equivalent to R. □

Proof of Theorem 1.5. Step 1. We prove (1-8) when R is an H-regular hypersurface. Then, R is locally the boundary of an open set $E \subset \mathbb{H}^n$ with H-regular boundary. Moreover, we have (locally) $\mu_E = \mathscr{S}^{2n+1} \llcorner R$ and $\nu_E = \nu_R$, up to the sign.

We define the measures $\mu_R^s = \mu_E^s$ for any s such that $\nabla u \neq 0$ on Σ^s. The measurability of the function in (1-7) follows from the argument (2-34). Formula (1-8) follows from Proposition 2.8.

Step 2. We prove (1-8) when R is an $\mathscr{S}^{2n+1}$-rectifiable Borel set. There exist an $\mathscr{S}^{2n+1}$-negligible set N and H-regular hypersurfaces $S_j \subset \mathbb{H}^n$, $j \in \mathbb{N}$, such that

$$R \subset N \cup \bigcup_{j=1}^{\infty} S_j.$$

Each S_j is (locally) the boundary of an open set E_j with H-regular boundary. We denote by $\mu_{E_j}^s$ the perimeter measure on $\partial E_j \cap \Sigma^s$ induced by E_j.

We define the pairwise disjoint Borel sets $R_j = (R \cap S_j) \setminus \bigcup_{h=1}^{j-1} S_h$ and we let

$$\mu_R^s = \sum_{j=1}^{\infty} \mu_{E_j}^s \llcorner R_j.$$

The definition is well posed for any s such that $\nabla u \neq 0$ on Σ^s. We have $\nu_R = \pm \nu_{E_j}$ $\mathscr{S}^{2n+1}$-a.e. on R_j and the sign of ν_R does not affect (1-8). From Step 1, for each $j \in \mathbb{N}$ the function

$$s \mapsto \int_{R_j} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_{E_j}^s$$

is $\mathscr{L}^1$-measurable; here, we were allowed to utilize Step 1 because χ_{R_j} is Borel regular. Thus also the function

$$s \mapsto \int_{\Omega} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_R^s = \sum_{j=1}^{\infty} \int_{R_j} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_{E_j}^s$$

is measurable. Moreover, we have

$$\begin{aligned}
\int_{\mathbb{R}} \int_{\Omega} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_R^s \, ds &= \sum_{j=1}^{\infty} \int_{\mathbb{R}} \int_{R_j} h \frac{|\nabla_H u|_g}{|\nabla u|_g} \, d\mu_{E_j}^s \, ds \\
&= \sum_{j=1}^{\infty} \int_{R_j} h \sqrt{|\nabla_H u|_g^2 - \langle \nu_R, \nabla_H u \rangle_g^2} \, d\mathscr{S}^{2n+1} \\
&= \int_R h \sqrt{|\nabla_H u|_g^2 - \langle \nu_R, \nabla_H u \rangle_g^2} \, d\mathscr{S}^{2n+1}.
\end{aligned}$$

Step 3. Finally, if R is $\mathscr{S}^{2n+1}$-rectifiable but not Borel, we set $\mu_R^s = \mu_{R'}^s$, where R' is a Borel set as in Lemma 2.9. Again, this definition is well posed for a.e. $s \in \mathbb{R}$. This concludes the proof. □

2C. *Proof of Theorem 1.6.* In this subsection we assume $n \geq 2$.

Lemma 2.10. *For $n \geq 2$, let $\Omega \subset \mathbb{H}^n$ be an open set, $u \in C^\infty(\Omega)$ a smooth function, $R \subset \Omega$ an $\mathscr{S}^{2n+1}$-rectifiable set. Then*

$$\mathscr{S}^{2n+1}\big(\{p \in R : \nabla_H u(p) = 0 \ \textit{and} \ \nabla u(p) \neq 0\}\big) = 0.$$

Proof. It is enough to prove the lemma when R is an H-regular hypersurface. Let

$$A = \{p \in R : \nabla_H u(p) = 0 \ \text{and} \ \nabla u(p) \neq 0\}.$$

We claim that $\mathscr{S}^{2n+1}(A) = 0$.

Let $p \in A$ be a fixed point and let $\nu_R(p)$ be the horizontal normal to R at p. Since $n \geq 2$, we have

$$\dim\{V(p) \in H_p : \langle V(p), \nu_R(p) \rangle_g = 0\} = 2n - 1 \geq n + 1.$$

Thus there exist left-invariant horizontal vector fields V and W such that

$$\langle V(p), \nu_R(p) \rangle_g = \langle W(p), \nu_R(p) \rangle_g = 0 \quad \text{and} \quad [V, W] = T.$$

From $\nabla_H u(p) = 0$ and $\nabla u(p) \neq 0$, we deduce that $Tu(p) \neq 0$. It follows that

$$VWu(p) - WVu(p) = Tu(p) \neq 0$$

and, in particular, we have either $VWu(p) \neq 0$ or $WVu(p) \neq 0$. Without loss of generality, we assume that $VWu(p) \neq 0$. Then the set $S = \{q \in \Omega : Wu(q) = 0\}$ is an H-regular hypersurface near the point $p \in S$. Since we have

$$\langle V(p), \nu_R(p)\rangle_g = 0 \quad \text{and} \quad \langle V(p), \nu_S(p)\rangle_g = \frac{VWu(p)}{|\nabla_H Wu(p)|_g} \neq 0,$$

we deduce that $\nu_R(p)$ and $\nu_S(p)$ are linearly independent. Then there exists $r > 0$ such that the set $R \cap S \cap B_r(p)$ is a 2-codimensional H-regular surface (see [Franchi et al. 2007]). Therefore, by [Franchi et al. 2007, Corollary 4.4], the Hausdorff dimension in the Carnot–Carathéodory metric of $A \cap B_r(p) \subset R \cap S \cap B_r(p)$ is not greater than $2n$. This is enough to conclude. □

Remark 2.11. Lemma 2.10 is not valid if $n = 1$. Consider the smooth surface $R = \{(x, y, t) \in \mathbb{H}^1 : x = 0\}$ and the function $u(x, y, t) = t - 2xy$. We have

$$\nabla u = -4xY + T \quad \text{and} \quad \nabla_H u = -4xY.$$

Then we have

$$\{p \in R : \nabla_H u(p) = 0 \text{ and } \nabla u(p) \neq 0\} = R$$

and $\mathscr{S}^3(R) = \infty$.

If $n \geq 2$ and Ω, u, and R are as in Lemma 2.10, then the function

$$|\nabla u|_g \sqrt{1 - \langle \nu_E, \nabla_H u / |\nabla_H u|_g\rangle_g^2}$$

is defined $\mathscr{S}^{2n+1}$-a.e. on R. We agree that its value is 0 when $|\nabla u|_g = 0$. Notice that, in this case, $\nabla_H u / |\nabla_H u|_g$ is not defined.

Proof of Theorem 1.6. Let $\varepsilon > 0$ be fixed. Then (1-9) can be obtained by plugging the function $(|\nabla u|_g / (\varepsilon + |\nabla_H u|_g))\, h$ into (1-8), letting $\varepsilon \to 0$ and using the monotone convergence theorem. □

3. Height estimate

In this section, we prove Theorem 1.3. We discuss first a relative isoperimetric inequality on slices. Then we list some elementary properties of excess, and finally we proceed with the proof.

We assume throughout this section that $n \geq 2$.

3A. *Relative isoperimetric inequalities.* For each $s \in \mathbb{R}$, we define the level sets of the height function,

$$\mathbb{H}^n_s = \{p \in \mathbb{H}^n : \mathscr{h}(p) = s\}.$$

Let H^s be the g-orthogonal projection of H onto the tangent space of $\mathbb{H}^n_s$. Since the vector field X_1 is orthogonal to $\mathbb{H}^n_s$, while the vector fields $X_2, \ldots, X_n, Y_1, \ldots, Y_n$ are tangent to $\mathbb{H}^n_s$, at any point $p \in \mathbb{H}^n_s$

we have

$$H_p^s = \operatorname{span}\{X_2(p), \dots, X_n(p), Y_1^s(p), Y_2(p), \dots, Y_n(p)\},$$

where $X_2, Y_2, \dots, X_n, Y_n$ are as in (1-2) and

$$Y_1^s = \frac{\partial}{\partial y_1} - 2s\frac{\partial}{\partial t}.$$

The natural volume in $\mathbb{H}_s^n$ is the Lebesgue measure $\mathscr{L}^{2n}$. For any measurable set $F \subset \mathbb{H}_s^n$ and any open set $\Omega \subset \mathbb{H}_s^n$, we define

$$\mu_F^s(\Omega) = \sup\left\{\int_F \operatorname{div}_g^s \varphi \, d\mathscr{L}^{2n} : \varphi \in C_c^1(\Omega; H^s), \|\varphi\|_g \le 1\right\},$$

where $\operatorname{div}_g^s \varphi = X_2\varphi_2 + \cdots + X_n\varphi_n + Y_1^s\varphi_{n+1} + \cdots + Y_n\varphi_{2n}$. If $\mu_F^s(\Omega) < \infty$ then μ_F^s is a Radon measure in Ω.

By Theorem 1.6, for any Borel function $h : \mathbb{H}^n \to [0, \infty)$ and any set E with locally finite H-perimeter in $\mathbb{H}^n$, we have the coarea formula

$$\int_{\mathbb{R}} \int_{\mathbb{H}_s^n} h \, d\mu_{E^s}^s \, ds = \int_{\mathbb{H}^n} h\sqrt{1 - \langle \nu_E, X_1\rangle_g^2} \, d\mu_E, \tag{3-41}$$

where $E^s = E \cap \mathbb{H}_s^n$ is the section of E with $\mathbb{H}_s^n$. Notice that $\nabla_H \mathscr{E} = X_1$.

In the proof of Theorem 1.3, we need a relative isoperimetric inequality in each slice $\mathbb{H}_s^n$ for $s \in (-1, 1)$. These slices are cosets of $\mathbb{W} = \mathbb{H}_0^n$ and the isoperimetric inequalities in $\mathbb{H}_s^n$ can be reduced to an isoperimetric inequality in the central slice $\mathbb{W} = \mathbb{H}_0^n$ relative to a family of varying domains.

For any $s \in (-1, 1)$, let $\Omega_s \subset \mathbb{W}$ be the set $\Omega_s = (-s\mathrm{e}_1) * D_1 * (s\mathrm{e}_1)$. This is the left translation by $-s\mathrm{e}_1$ of the section $C_1 \cap \mathbb{H}_s^n$. See p. 1423 in the introduction for the definition of D_1 and C_1. With the coordinates $(y_1, \hat{z}, t) \in \mathbb{W} = \mathbb{R} \times \mathbb{C}^{n-1} \times \mathbb{R}$, we have

$$\Omega_s = \{(y_1, \hat{z}, t) \in \mathbb{W} : (y_1^2 + |\hat{z}|^2)^2 + (t - 4sy_1)^2 < 1\}.$$

The sets $\Omega_s \subset \mathbb{W}$ are open and convex in the standard sense. The boundary $\partial\Omega_s$ is a $(2n-1)$-dimensional C^∞ embedded surface with the following property: There are $4n$ open convex sets $U_1, \dots, U_{4n} \subset \mathbb{W}$ such that $\partial\Omega_s \subset \bigcup_{i=1}^{4n} U_i$ and, for each i, the portion of the boundary $\partial\Omega_s \cap U_i$ is a graph of the form $p_j = f_i^s(\hat{p}_j)$ with $j = 2, \dots, 2n+1$ and $\hat{p}_j = (p_2, \dots, p_{j-1}, p_{j+1}, \dots, p_{2n+1}) \in V_i$, where $V_i \subset \mathbb{R}^{2n-1}$ is an open convex set and $f_i^s \in C^\infty(V_i)$ is a function such that

$$|\nabla f_i^s(\hat{p}_j) - \nabla f_i^s(\hat{q}_j)| \le K|\hat{p}_j - \hat{q}_j| \quad \text{for all } \hat{p}_j, \hat{q}_j \in V_i, \tag{3-42}$$

where $K > 0$ is a constant independent of $i = 1, \dots, 4n$ and independent of $s \in (-1, 1)$. In other words, the boundary $\partial\Omega_s$ is of class $C^{1,1}$ uniformly in $s \in (-1, 1)$.

By Theorem 3.2 in [Monti and Morbidelli 2005], the domain $\Omega_s \subset \mathbb{W}$ is a nontangentially accessible (NTA) domain in the metric space $(\mathbb{W}, d_{CC})$, where d_{CC} is the Carnot–Carathéodory metric induced by the horizontal distribution H_p^0. In particular, Ω_s is a (weak) John domain in the sense of [Hajłasz and

Koskela 2000]. Namely, there exist a point $p_0 \in \Omega_s$, e.g., $p_0 = 0$, and a constant $C_J > 0$ such that, for any point $p \in \Omega_s$, there exists a continuous curve $\gamma : [0, 1] \to \Omega_s$ such that $\gamma(1) = p_0$, $\gamma(0) = p$, and

$$d_{CC}(\gamma(\sigma), \partial\Omega_s) \geq C_J d_{CC}(\gamma(\sigma), p), \quad \sigma \in [0, 1]. \tag{3-43}$$

By Theorem 3.2 in [Monti and Morbidelli 2005], the John constant C_J depends only on the constant $K > 0$ in (3-42). This claim is not stated explicitly in Theorem 3.2 of [Monti and Morbidelli 2005] but it is evident from the proof. In particular, the John constant C_J is independent of $s \in (-1, 1)$. Then, by Theorem 1.22 in [Garofalo and Nhieu 1996], we have the following result:

Theorem 3.1. *Let $n \geq 2$. There exists a constant $C(n) > 0$ such that, for any $s \in (-1, 1)$ and any measurable set $F \subset \mathbb{W}$,*

$$\min\{\mathscr{L}^{2n}(F \cap \Omega_s), \mathscr{L}^{2n}(\Omega_s \setminus F)\}^{2n/(2n+1)} \leq C(n) \frac{\operatorname{diam}_{CC}(\Omega_s)}{\mathscr{L}^{2n}(\Omega_s)^{1/(2n+1)}} \mu_F^0(\Omega_s). \tag{3-44}$$

An alternative proof of Theorem 3.1 can be obtained using the Sobolev–Poincaré inequalities proved in [Hajłasz and Koskela 2000] in the general setting of metric spaces.

The diameter $\operatorname{diam}_{CC}(\Omega_s)$ is bounded for $s \in (-1, 1)$ and $\mathscr{L}^{2n}(\Omega_s) > 0$ is a constant independent of s. Then we obtain the following version of (3-44):

Corollary 3.2. *Let $n \geq 2$. For any $\tau \in (0, 1)$ there exists a constant $C(n, \tau) > 0$ such that, for $s \in (-1, 1)$ and any measurable set $F \subset \mathbb{W}$ satisfying*

$$\mathscr{L}^{2n}(F \cap \Omega_s) \leq \tau \mathscr{L}^{2n}(\Omega_s),$$

we have

$$\mu_F^0(\Omega_s) \geq C(n, \tau) \mathscr{L}^{2n}(F \cap \Omega_s)^{2n/(2n+1)}.$$

3B. ***Elementary properties of the excess.*** We list here, without proof, the most basic properties of the cylindrical excess introduced in Definition 1.2. Their proofs are easy adaptations of those for the classical excess; see, e.g., [Maggi 2012, Chapter 22]. Note that, except for property (3), they hold also in the case $n = 1$.

(1) For all $0 < r < s$, we have

$$\operatorname{Exc}(E, r, \nu) \leq \left(\frac{s}{r}\right)^{2n+1} \operatorname{Exc}(E, s, \nu). \tag{3-45}$$

(2) If $(E_j)_{j\in\mathbb{N}}$ is a sequence of sets with locally finite H-perimeter such that $E_j \to E$ as $j \to \infty$ in $L^1_{\text{loc}}(\mathbb{H}^n)$, then we have, for any $r > 0$,

$$\operatorname{Exc}(E, r, \nu) \leq \liminf_{j\to\infty} \operatorname{Exc}(E_j, r, \nu). \tag{3-46}$$

(3) Let $n \geq 2$. If $E \subset \mathbb{H}^n$ is a set such that $\operatorname{Exc}(E, r, \nu) = 0$ and $0 \in \partial^* E$, then

$$E \cap C_r = \{p \in C_r : \mathscr{E}(p) < 0\}. \tag{3-47}$$

In particular, we have $\nu_E = \nu$ in $C_r \cap \partial E$. See also [Monti 2014, Proposition 3.6].

(4) For any $\lambda > 0$ and $r > 0$, we have
$$\operatorname{Exc}(\lambda E, \lambda r, \nu) = \operatorname{Exc}(E, r, \nu), \tag{3-48}$$
where $\lambda E = \{(\lambda z, \lambda^2 t) \in \mathbb{H}^n : (z, t) \in E\}$.

3C. *Proof of Theorem 1.3.* The following result is a first, suboptimal version of Theorem 1.3.

Lemma 3.3. *Let $n \geq 2$. For any $s \in (0, 1)$, $\Lambda \in [0, \infty)$, and $r \in (0, \infty]$ with $\Lambda r \leq 1$, there exists a constant $\omega(n, s, \Lambda, r) > 0$ such that, if $E \subset \mathbb{H}^n$ is a (Λ, r)-minimizer of H-perimeter in the cylinder C_2, $0 \in \partial E$, and $\operatorname{Exc}(E, 2, \nu) \leq \omega(n, s, \Lambda, r)$, then*
$$\begin{gathered} |\mathfrak{k}(p)| < s \quad \textit{for any } p \in \partial E \cap C_1, \\ \mathscr{L}^{2n+1}\big(\{p \in E \cap C_1 : \mathfrak{k}(p) > s\}\big) = 0, \\ \mathscr{L}^{2n+1}\big(\{p \in C_1 \setminus E : \mathfrak{k}(p) < -s\}\big) = 0. \end{gathered}$$

Proof. By contradiction, assume that there exist $s \in (0, 1)$ and a sequence of sets $(E_j)_{j\in\mathbb{N}}$ that are (Λ, r)-minimizers in C_2 and such that
$$\lim_{j\to\infty} \operatorname{Exc}(E_j, 2, \nu) = 0$$
and at least one of the following facts holds:
$$\text{there exists } p \in \partial E_j \cap C_1 \text{ such that } s \leq |\mathfrak{k}(p)| \leq 1, \tag{3-49}$$
$$\mathscr{L}^{2n+1}\big(\{p \in E_j \cap C_1 : \mathfrak{k}(p) > s\}\big) > 0, \tag{3-50}$$
$$\text{or} \quad \mathscr{L}^{2n+1}\big(\{p \in C_1 \setminus E_j : \mathfrak{k}(p) < -s\}\big) > 0. \tag{3-51}$$
By Theorem A.3 in Appendix A, there exists a measurable set $F \subset C_{5/3}$ such that F is a (Λ, r)-minimizer in $C_{5/3}$, $0 \in \partial F$, and (possibly up to subsequences) $E_j \cap C_{5/3} \to F$ in $L^1(C_{5/3})$. By (3-46) and (3-45), we obtain
$$\operatorname{Exc}\big(F, \tfrac{4}{3}, \nu\big) \leq \liminf_{j\to\infty} \operatorname{Exc}\big(E_j, \tfrac{4}{3}, \nu\big) \leq \big(\tfrac{3}{2}\big)^{2n+1} \lim_{j\to\infty} \operatorname{Exc}(E_j, 2, \nu) = 0.$$
Since $0 \in \partial F$, by (3-47) the set $F \cap C_{4/3}$ is (equivalent to) a halfspace with horizontal inner normal $\nu = -X_1$, namely,
$$F \cap C_{4/3} = \{p \in C_{4/3} : \mathfrak{k}(p) < 0\}.$$
Assume that (3-49) holds for infinitely many j. Then, up to a subsequence, there are points $(p_j)_{j\in\mathbb{N}}$ and p_0 such that
$$p_j \in \partial E_j \cap C_1, \quad |\mathfrak{k}(p_j)| \in (s, 1] \quad \text{and} \quad p_j \to p_0 \in \partial F \cap \bar{C}_1.$$
We used again Theorem A.3 in Appendix A. This is a contradiction because $\partial F \cap \bar{C}_1 = \{p \in \bar{C}_1 : \mathfrak{k}(p) = 0\}$. Here, we used $n \geq 2$. Therefore, there exists $j_0 \in \mathbb{N}$ such that
$$\{p \in \partial E_j \cap C_1 : s \leq |\mathfrak{k}(p)| \leq 1\} = \varnothing \quad \text{for all } j \geq j_0$$

and hence

$$\mu_{E_j}\big(C_1 \setminus \{p \in \mathbb{H}^n : |\not{z}(p)| \le s\}\big) = 0.$$

This implies that, for $j \ge j_0$, χ_{E_j} is constant on the two connected components $C_1 \cap \{p : \not{z}(p) > s\}$ and $C_1 \cap \{p : \not{z}(p) < -s\}$. Since the sequence $(E_j)_{j\in\mathbb{N}}$ converges in $L^1(C_1)$ to the halfspace F, for any $j \ge j_0$ we have

$$\begin{aligned} &\chi_{E_j} = 0 \quad \mathscr{L}^{2n+1}\text{-a.e. on } C_1 \cap \{p : \not{z}(p) > s\}, \\ \text{and} \qquad &\chi_{E_j} = 1 \quad \mathscr{L}^{2n+1}\text{-a.e. on } C_1 \cap \{p : \not{z}(p) < -s\}. \end{aligned}$$

This contradicts both (3-50) and (3-51) and concludes the proof. □

Let $\pi : \mathbb{H}^n \to \mathbb{W}$ be the group projection defined, for any $p \in \mathbb{H}^n$, by the formula

$$p = \pi(p) * (\not{z}(p)\mathrm{e}_1).$$

For any set $E \subset \mathbb{H}^n$ and $s \in \mathbb{R}$, we let $E^s = E \cap \mathbb{H}^n_s$ and we define the projection

$$E_s = \pi(E^s) = \{w \in \mathbb{W} : w * (s\mathrm{e}_1) \in E\}.$$

Lemma 3.4. *Let $n \ge 2$, let $E \subset \mathbb{H}^n$ be a set with locally finite H-perimeter and $0 \in \partial E$, and let $s_0 \in (0, 1)$ be such that*

$$|\not{z}(p)| < s_0 \quad \textit{for any } p \in \partial E \cap C_1, \tag{3-52}$$

$$\mathscr{L}^{2n+1}\big(\{p \in E \cap C_1 : \not{z}(p) > s_0\}\big) = 0, \tag{3-53}$$

$$\mathscr{L}^{2n+1}\big(\{p \in C_1 \setminus E : \not{z}(p) < -s_0\}\big) = 0. \tag{3-54}$$

Then, for a.e. $s \in (-1, 1)$ and any continuous function $\varphi \in C_c(D_1)$, we have, with $M = \partial^* E \cap C_1$ and $M_s = M \cap \{\not{z} > s\}$,

$$\int_{E_s \cap D_1} \varphi \, d\mathscr{L}^{2n} = -\int_{M_s} \varphi \circ \pi \, \langle \nu_E, X_1 \rangle_g \, d\mathscr{S}^{2n+1}. \tag{3-55}$$

In particular, for any Borel set $G \subset D_1$, we have

$$\mathscr{L}^{2n}(G) = -\int_{M \cap \pi^{-1}(G)} \langle \nu_E, X_1 \rangle_g \, d\mathscr{S}^{2n+1}, \tag{3-56}$$

$$\mathscr{L}^{2n}(G) \le \mathscr{S}^{2n+1}(M \cap \pi^{-1}(G)). \tag{3-57}$$

Proof. It is enough to prove (3-55). Indeed, taking $s < -s_0$ in (3-55) and recalling (3-52) and (3-54), we obtain

$$\int_{D_1} \varphi \, d\mathscr{L}^{2n} = -\int_M \varphi \circ \pi \, \langle \nu_E, X_1 \rangle_g \, d\mathscr{S}^{2n+1}. \tag{3-58}$$

Formula (3-56) follows from (3-58) by considering smooth approximations of χ_G. Formula (3-57) is immediate from (3-56) and $|\langle \nu_E, X_1 \rangle_g| \le 1$.

We prove (3-55) for a.e. $s \in (-1, 1)$, namely, for those s satisfying the property (3-61) below. Up to an approximation argument, we may assume that $\varphi \in C_c^1(D_1)$. Let $r \in (0, 1)$ and $\sigma \in (\max\{s_0, s\}, 1)$ be fixed. We define

$$F = E \cap (D_r * (s, \sigma)) = E \cap \{w * (\varrho \mathrm{e}_1) \in \mathbb{H}^n : w \in D_r,\ \varrho \in (s, \sigma)\}.$$

We claim that, for a.e. $r \in (0, 1)$ and any s satisfying (3-61), we have

$$\langle \nu_F, X_1\rangle_g \mu_F = \langle \nu_E, X_1\rangle_g \mathscr{S}^{2n+1} \llcorner \partial^* E \cap (D_r * (s, \sigma)) + \mathscr{L}^{2n} \llcorner E \cap D_r^s. \tag{3-59}$$

Above, we let $D_r^s = \{w * (s\mathrm{e}_1) \in \mathbb{H}^n : w \in D_r\}$. We postpone the proof of (3-59). Let Z be a horizontal vector field of the form $Z = (\varphi \circ \pi)X_1$. We have $\operatorname{div}_g Z = 0$ because $X_1(\varphi \circ \pi) = 0$. Hence, we obtain

$$0 = \int_F \operatorname{div}_g Z \, d\mathscr{L}^{2n+1} = -\int_{\mathbb{H}^n} \varphi \circ \pi \, \langle \nu_F, X_1\rangle_g \, d\mu_F,$$

i.e., by the Fubini–Tonelli theorem and (3-59),

$$-\int_{E_s \cap D_r} \varphi \, d\mathscr{L}^{2n} = -\int_{E \cap D_r^s} \varphi \circ \pi \, d\mathscr{L}^{2n} = \int_{\partial^* E \cap (D_r * (s,\sigma))} \varphi \circ \pi \, \langle \nu_E, X_1\rangle_g \, d\mathscr{S}^{2n+1}.$$

Formula (3-55) follows on letting first $r \nearrow 1$ and then $\sigma \nearrow 1$.

We are left with the proof of (3-59). Let $\psi \in C_c^1(\mathbb{H}^n)$ be a test function. For any $w \in \mathbb{W}$, we let

$$E_w = \{\varrho \in \mathbb{R} : w * (\varrho \mathrm{e}_1) \in E\}, \qquad \psi_w(\varrho) = \psi(w * (\varrho \mathrm{e}_1)).$$

Then we have $\psi_w \in C_c^1(\mathbb{R})$ and, by the Fubini–Tonelli theorem,

$$\begin{aligned}
-\int_F X_1 \psi \, d\mathscr{L}^{2n+1} &= -\int_{D_r} \int_s^\sigma \chi_E(w * (\varrho \mathrm{e}_1)) X_1 \psi(w * (\varrho \mathrm{e}_1)) \, d\varrho \, d\mathscr{L}^{2n}(w) \\
&= -\int_{D_r} \int_s^\sigma \chi_{E_w}(\varrho) \psi_w'(\varrho) \, d\varrho \, d\mathscr{L}^{2n}(w) \\
&= \int_{D_r} \left[\int_s^\sigma \psi_w \, dD\chi_{E_w} - \psi_w(\sigma)\chi_{E_w}(\sigma^-) + \psi_w(s)\chi_{E_w}(s^+) \right] d\mathscr{L}^{2n}(w),
\end{aligned} \tag{3-60}$$

where $D\chi_{E_w}$ is the derivative of χ_{E_w} in the sense of distributions and $\chi_{E_w}(\sigma^-)$, $\chi_{E_w}(s^+)$ are the classical trace values of χ_{E_w} at the endpoints of the interval (s, σ). We used the fact that the function χ_{E_w} is of bounded variation for $\mathscr{L}^{2n}$-a.e. $w \in \mathbb{W}$, which in turn is a consequence of the fact that $X_1 \chi_E$ is a signed Radon measure. For any such w, the trace of χ_{E_w} satisfies

$$\chi_{E_w}(s^+) = \chi_{E_w}(s) = \chi_E(w * (s\mathrm{e}_1)) \quad \text{for a.e. } s,$$

so that, by Fubini's theorem, for a.e. $s \in \mathbb{R}$ we have

$$\chi_{E_w}(s^+) = \chi_E(w * (s\mathrm{e}_1)) \quad \text{for } \mathscr{L}^{2n}\text{-a.e. } w \in D_1. \tag{3-61}$$

With a similar argument, using (3-53) and the fact that $\sigma > s_0$, one can see that

$$\chi_{E_w}(\sigma^-) = \chi_E(w * (\sigma \mathrm{e}_1)) = 0 \quad \text{for } \mathscr{L}^{2n}\text{-a.e. } w \in D_1. \tag{3-62}$$

We refer the reader to [Ambrosio et al. 2000] for an extensive account on BV functions and traces. By (3-60), (3-61) and (3-62), we obtain

$$\begin{aligned}
-\int_F X_1\psi\, d\mathscr{L}^{2n+1} &= \int_{D_r}\int_s^\sigma \psi_w\, dD\chi_{E_w}\, d\mathscr{L}^{2n}(w) + \int_{D_r} \psi_w(s)\chi_{E_w}(s)\, d\mathscr{L}^{2n}(w)\\
&= \int_{D_r*(s,\sigma)} \psi\, \langle \nu_E, X_1\rangle_g d\mu_E + \int_{E\cap D_r^s} \psi\, d\mathscr{L}^{2n}\\
&= \int_{\partial^*E\cap(D_r*(s,\sigma))} \psi\, \langle \nu_E, X_1\rangle_g\, d\mathscr{S}^{2n+1} + \int_{E\cap D_r^s} \psi\, d\mathscr{L}^{2n},
\end{aligned}$$

and (3-59) follows. □

Corollary 3.5. *Under the same assumptions and notation as Lemma 3.4, for a.e. $s \in (-1, 1)$, we have*

$$0 \le \mathscr{S}^{2n+1}(M_s) - \mathscr{L}^{2n}(E_s \cap D_1) \le \operatorname{Exc}(E, 1, \nu). \tag{3-63}$$

Moreover,

$$\mathscr{S}^{2n+1}(M) - \mathscr{L}^{2n}(D_1) = \operatorname{Exc}(E, 1, \nu). \tag{3-64}$$

Proof. On approximating χ_{D_1} with functions $\varphi \in C_c(D_1)$, by (3-55) we get

$$\mathscr{L}^{2n}(E_s \cap D_1) = -\int_{M_s} \langle \nu_E, X_1\rangle_g\, d\mathscr{S}^{2n+1},$$

and the first inequality in (3-63) follows. The second inequality follows from

$$\begin{aligned}
\mathscr{S}^{2n+1}(M_s) - \mathscr{L}^{2n}(E_s \cap D_1) &= \int_{M_s} (1 + \langle \nu_E, X_1\rangle_g)\, d\mathscr{S}^{2n+1}\\
&= \int_{M_s} \frac{|\nu_E - \nu|_g^2}{2}\, d\mathscr{S}^{2n+1}\\
&\le \operatorname{Exc}(E, 1, \nu).
\end{aligned} \tag{3-65}$$

Notice that $\nu = -X_1$. Finally, (3-64) follows on choosing a suitable $s < -s_0$ and recalling (3-52) and (3-54). In this case, the inequality in (3-65) becomes an equality and the proof is concluded. □

Proof of Theorem 1.3. Step 1. Up to replacing E with the rescaled set $\lambda E = \{(\lambda z, \lambda^2 t) \in \mathbb{H}^n : (z, t) \in E\}$ with $\lambda = 1/2k^2 r$ and recalling (3-48), we can without loss of generality assume that E is a $(\Lambda', 1/(2k^2))$-minimizer of H-perimeter in C_2 with

$$\frac{\Lambda'}{2k^2} \le 1, \quad 0 \in \partial E, \quad \operatorname{Exc}(E, 2, \nu) \le \varepsilon_0(n). \tag{3-66}$$

Our goal is to find $\varepsilon_0(n)$ and $c_1(n) > 0$ such that, if (3-66) holds, then

$$\sup\{|\mathfrak{k}(p)| : p \in \partial E \cap C_{1/2k^2}\} \le c_1(n)\operatorname{Exc}(E, 2, \nu)^{1/(2(2n+1))}. \tag{3-67}$$

We require

$$\varepsilon_0(n) \le \omega\Big(n, \frac{1}{4k}, 2k^2, \frac{1}{2k^2}\Big), \tag{3-68}$$

where ω is as given by Lemma 3.3. Two further assumptions on $\varepsilon_0(n)$ will be made later, in (3-80) and (3-85). By (3-66), E is a $(2k^2, 1/(2k^2))$-minimizer in C_2. Letting $M = \partial E \cap C_1$, by Lemma 3.3 and (3-68) we have

$$|\mathfrak{k}(p)| < \frac{1}{4k} \quad \text{for any } \ p \in M, \tag{3-69}$$

$$\mathscr{L}^{2n+1}\left(\left\{p \in E \cap C_1 : \mathfrak{k}(p) > \frac{1}{4k}\right\}\right) = 0, \tag{3-70}$$

$$\mathscr{L}^{2n+1}\left(\left\{p \in C_1 \setminus E : \mathfrak{k}(p) < -\frac{1}{4k}\right\}\right) = 0. \tag{3-71}$$

By (3-64) and (3-45), we get

$$0 \le \mathscr{S}^{2n+1}(M) - \mathscr{L}^{2n}(D_1) \le \operatorname{Exc}(E, 1, \nu) \le 2^{2n+1} \operatorname{Exc}(E, 2, \nu). \tag{3-72}$$

Corollary 3.5 implies that, for a.e. $s \in (-1, 1)$,

$$0 \le \mathscr{S}^{2n+1}(M_s) - \mathscr{L}^{2n}(E_s \cap D_1) \le \operatorname{Exc}(E, 1, \nu) \le 2^{2n+1} \operatorname{Exc}(E, 2, \nu), \tag{3-73}$$

where, as before, $M_s = M \cap \{\mathfrak{k} > s\}$.

Step 2. Consider $f : (-1, 1) \to [0, \mathscr{S}^{2n+1}(M)]$ defined by

$$f(s) = \mathscr{S}^{2n+1}(M_s), \qquad s \in (-1, 1).$$

The function f is nonincreasing, right-continuous and, by (3-69), it satisfies

$$\begin{aligned} f(s) &= \mathscr{S}^{2n+1}(M) && \text{for any } \ s \in \left(-1, -\frac{1}{4k}\right], \\ f(s) &= 0 && \text{for any } \ s \in \left(\frac{1}{4k}, 1\right]. \end{aligned}$$

In particular, there exists $s_0 \in (-1/(4k), 1/(4k))$ such that

$$\begin{aligned} f(s) &\ge \tfrac{1}{2}\mathscr{S}^{2n+1}(M) && \text{for any } \ s < s_0, \\ f(s) &\le \tfrac{1}{2}\mathscr{S}^{2n+1}(M) && \text{for any } \ s \ge s_0. \end{aligned} \tag{3-74}$$

Let $s_1 \in (s_0, 1/(4k))$ be such that

$$\begin{aligned} f(s) &\ge \sqrt{\operatorname{Exc}(E, 2, \nu)} && \text{for any } \ s < s_1, \\ f(s) &= \mathscr{S}^{2n+1}(M_s) \le \sqrt{\operatorname{Exc}(E, 2, \nu)} && \text{for any } \ s \ge s_1. \end{aligned} \tag{3-75}$$

We claim that there exists $c_2(n) > 0$ such that

$$\mathfrak{k}(p) \le s_1 + c_2(n) \operatorname{Exc}(E, 2, \nu)^{1/(2(2n+1))} \quad \text{for any } \ p \in \partial E \cap C_{1/2k^2}. \tag{3-76}$$

The inequality (3-76) is trivial for any $p \in \partial E \cap C_{1/2k^2}$ with $\mathfrak{k}(p) \le s_1$. If $p \in \partial E \cap C_{1/2k^2}$ is such that $\mathfrak{k}(p) > s_1$, then

$$B_{\mathfrak{k}(p)-s_1}(p) \subset B_{1/2k}(p) \subset B_{1/k} \subset C_1.$$

We used the fact that $\|p\|_K \le 1/(2k)$ whenever $p \in C_{1/2k^2}$; see (1-3). Therefore,

$$B_{\mathscr{E}(p)-s_1}(p) \subset C_1 \cap \{\mathscr{E} > s_1\}$$

and, by the density estimate (A-91) of Theorem A.1 in Appendix A,

$$k_3(n)(\mathscr{E}(p)-s_1)^{2n+1} \le \mu_E(B_{\mathscr{E}(p)-s_1}(p)) \le \mu_E(C_1 \cap \{\mathscr{E} > s_1\}) = \mathscr{S}^{2n+1}(M_{s_1}) = f(s_1) \le \sqrt{\operatorname{Exc}(E, 2, \nu)}.$$

This proves (3-76).

Step 3. We claim that there exists $c_3(n) > 0$ such that

$$s_1 - s_0 \le c_3(n) \operatorname{Exc}(E, 2, \nu)^{1/(2(2n+1))}. \tag{3-77}$$

By the coarea formula (3-41) with $h = \chi_{C_1}$, $D_1^s = \{p \in C_1 : \mathscr{E}(p) = s\}$, and $E^s = \{p \in E : \mathscr{E}(p) = s\}$, we have

$$\int_{-1}^{1} \int_{D_1^s} d\mu^s_{E^s}\, ds = \int_{C_1} \sqrt{1 - \langle \nu_E, X_1\rangle_g^2}\, d\mu_E \le \sqrt{2} \int_M \sqrt{1 + \langle \nu_E, X_1\rangle_g}\, d\mathscr{S}^{2n+1}.$$

By Hölder's inequality, (A-91), (3-56), and (3-72), we deduce that

$$\begin{aligned} \int_{-1}^{1} \int_{D_1^s} d\mu^s_{E^s}\, ds &\le \sqrt{2\mathscr{S}^{2n+1}(M)} \left(\int_M (1 + \langle \nu_E, X_1\rangle_g)\, d\mathscr{S}^{2n+1} \right)^{\frac{1}{2}} \\ &\le c_4(n)(\mathscr{S}^{2n+1}(M) - \mathscr{L}^{2n}(D_1))^{1/2} \\ &\le c_5(n) \sqrt{\operatorname{Exc}(E, 2, \nu)}. \end{aligned} \tag{3-78}$$

By Corollary 3.5 and (3-72), we obtain, for a.e. $s \in [s_0, s_1)$,

$$\begin{aligned} \mathscr{L}^{2n}(E_s \cap D_1) \le \mathscr{S}^{2n+1}(M_s) = f(s) \le f(s_0) &\le \tfrac{1}{2} \mathscr{S}^{2n+1}(M) \\ &\le \tfrac{1}{2}(\mathscr{L}^{2n}(D_1) + 2^{2n+1} \operatorname{Exc}(E, 2, \nu)) \\ &\le \tfrac{3}{4} \mathscr{L}^{2n}(D_1). \end{aligned} \tag{3-79}$$

The last inequality holds provided that

$$2^{2n+1} \varepsilon_0(n) \le \tfrac{1}{4} \mathscr{L}^{2n}(D_1). \tag{3-80}$$

Let $\Omega_s = (-se_1) * D_1^s = (-se_1) * D_1 * (se_1)$ and $F_s = (-se_1) * E^s$. We have

$$\mathscr{L}^{2n}(\Omega_s) = \mathscr{L}^{2n}(D_1^s) = \mathscr{L}^{2n}(D_1) \tag{3-81}$$

and, by (3-79),

$$\mathscr{L}^{2n}(F_s \cap \Omega_s) = \mathscr{L}^{2n}(E^s \cap D_1^s) = \mathscr{L}^{2n}(E_s \cap D_1) \le \tfrac{3}{4} \mathscr{L}^{2n}(D_1). \tag{3-82}$$

Moreover, by left invariance we have

$$\mu^s_{E^s}(D_1^s) = \mu^0_{F_s}(\Omega_s). \tag{3-83}$$

By (3-81)–(3-83) and Corollary 3.2, there exists a constant $k(n) > 0$ independent of $s \in (-1, 1)$ such that

$$\mu_{E^s}(D_1^s) = \mu^0_{F_s}(\Omega_s) \geq k(n)\mathscr{L}^{2n}(F_s \cap \Omega_s)^{2n/(2n+1)} = k(n)\mathscr{L}^{2n}(E^s \cap D_1^s)^{2n/(2n+1)}. \tag{3-84}$$

This, together with (3-78), gives

$$\begin{aligned} c_6(n)\sqrt{\operatorname{Exc}(E, 2, \nu)} &\geq \int_{s_0}^{s_1} \mathscr{L}^{2n}(E^s \cap D_1^s)^{2n/(2n+1)}\, ds \\ &\overset{(3\text{-}73)}{\geq} \int_{s_0}^{s_1} (\mathscr{S}^{2n+1}(M_s) - 2^{2n+1}\operatorname{Exc}(E, 2, \nu))^{2n/(2n+1)}\, ds \\ &\overset{(3\text{-}75)}{\geq} \int_{s_0}^{s_1} \big(\sqrt{\operatorname{Exc}(E, 2, \nu)} - 2^{2n+1}\operatorname{Exc}(E, 2, \nu)\big)^{2n/(2n+1)}\, ds \\ &\geq \frac{1}{2}\int_{s_0}^{s_1} \operatorname{Exc}(E, 2, \nu)^{n/(2n+1)}\, ds. \end{aligned}$$

In the last inequality, we require that $\varepsilon_0(n)$ satisfies

$$\sqrt{z} - 2^{2n+1}z \geq \tfrac{1}{2}\sqrt{z} \quad \text{for all } z \in [0, \varepsilon_0(n)]. \tag{3-85}$$

It follows that

$$c_6(n)\sqrt{\operatorname{Exc}(E, 2, \nu)} \geq \tfrac{1}{2}\operatorname{Exc}(E, 2, \nu)^{n/(2n+1)}(s_1 - s_0),$$

giving (3-77).

Step 4. Recalling (3-76) and (3-77), we proved that there exist $\varepsilon_0(n)$ and $c_6(n)$ such that the following holds: if E is a $(2k^2, 1/(2k^2))$-minimizer of H-perimeter in C_2 such that

$$0 \in \partial E, \quad \operatorname{Exc}(E, 2, \nu) \leq \varepsilon_0(n)$$

and $s_0 = s_0(E)$ satisfies (3-74), then

$$\mathscr{E}(p) - s_0 \leq c_7(n)\operatorname{Exc}(E, 2, \nu)^{1/(2(2n+1))} \quad \text{for any } p \in \partial E \cap C_{1/2k^2}. \tag{3-86}$$

Let us introduce the mapping $\Psi : \mathbb{H}^n \to \mathbb{H}^n$

$$\Psi(x_1, x_2 \ldots, x_n, y_1, \ldots, y_n, t) = (-x_1, -x_2, \ldots, -x_n, y_1, \ldots, y_n, -t)\,.$$

Then we have $\Psi^{-1} = \Psi$, $\Psi(C_2) = C_2$, $\langle X_j, \nu_{\Psi(F)}\rangle_g = -\langle X_j, \nu_F\rangle_g \circ \Psi$, $\langle Y_j, \nu_{\Psi(F)}\rangle_g = \langle Y_j, \nu_F\rangle_g \circ \Psi$, and $\mu_{\Psi(F)} = \Psi_\# \mu_F$, for any set F with locally finite H-perimeter; here, $\Psi_\#$ denotes the standard push-forward of measures. Therefore, the set $\widetilde{E} = \Psi(\mathbb{H}^n \setminus E)$ satisfies the following properties:

(i) $\widetilde{E}$ is a $(2k^2, 1/(2k^2))$-minimizer of H-perimeter in C_2;

(ii) $0 \in \partial\widetilde{E}$ and

$$\operatorname{Exc}(\widetilde{E}, 2, \nu) = \frac{1}{2^Q}\int_{\partial^*\widetilde{E}\cap C_2} |\nu_{\widetilde{E}} - \nu|_g^2\, d\mathscr{S}^{2n+1} = \operatorname{Exc}(E, 2, \nu) \leq \varepsilon_0(n);$$

(iii) setting $\widetilde{M} = \partial^*\widetilde{E} \cap C_1 = \Psi(M)$ and $\tilde{f}(s) = \mathscr{S}^{2n+1}(\widetilde{M} \cap \{\ell > s\})$, we have

$$\tilde{f}(s) \geq \tfrac{1}{2}\mathscr{S}^{2n+1}(\widetilde{M}) = \tfrac{1}{2}\mathscr{S}^{2n+1}(M) \quad \text{for any } s < -s_0,$$
$$\tilde{f}(s) \leq \tfrac{1}{2}\mathscr{S}^{2n+1}(M) \quad \text{for any } s \geq -s_0.$$

Formula (3-86) for the set $\widetilde{E}$ gives

$$\ell(p) + s_0 \leq c_7(n)\,\mathrm{Exc}(E, 2, \nu)^{1/(2(2n+1))} \quad \text{for any } p \in \partial\widetilde{E} \cap C_{1/2k^2}.$$

Notice that we have $p \in \partial\widetilde{E}$ if and only if $\Psi(p) \in \partial E$ and, moreover, $\ell(\Psi(p)) = -\ell(p)$. Hence, we have

$$-\ell(p) + s_0 \leq c_7(n)\,\mathrm{Exc}(E, 2, \nu)^{1/(2(2n+1))} \quad \text{for any } p \in \partial E \cap C_{1/2k^2}. \tag{3-87}$$

By (3-86) and (3-87), we obtain

$$|\ell(p) - s_0| \leq c_7(n)\,\mathrm{Exc}(E, 2, \nu)^{1/(2(2n+1))} \quad \text{for any } p \in \partial E \cap C_{1/2k^2}, \tag{3-88}$$

and, in particular,

$$|s_0| \leq c_7(n)\,\mathrm{Exc}(E, 2, \nu)^{1/(2(2n+1))}, \tag{3-89}$$

because $0 \in \partial E \cap C_{1/2k^2}$. Inequalities (3-88) and (3-89) give (3-67). This completes the proof. □

Appendix A

We list some basic properties of Λ-minimizers of H-perimeter in $\mathbb{H}^n$. The proofs are straightforward adaptations of the proofs for Λ-minimizers of perimeter in $\mathbb{R}^n$.

Theorem A.1 (density estimates). *There exist positive constants $k_1(n)$, $k_2(n)$, $k_3(n)$ and $k_4(n)$ with the following property: if E is a (Λ, r)-minimizer of H-perimeter in $\Omega \subset \mathbb{H}^n$, $p \in \partial E \cap \Omega$, $B_r(p) \subset \Omega$ and $s < r$, then*

$$k_1(n) \leq \frac{\mathscr{L}^{2n+1}(E \cap B_s(p))}{s^{2n+2}} \leq k_2(n), \tag{A-90}$$

$$k_3(n) \leq \frac{\mu_E(B_s(p))}{s^{2n+1}} \leq k_4(n). \tag{A-91}$$

For a proof, see [Maggi 2012, Theorem 21.11]. By standard arguments, Theorem A.1 implies the following corollary:

Corollary A.2. *If E is a (Λ, r)-minimizer of H-perimeter in $\Omega \subset \mathbb{H}^n$, then*

$$\mathscr{S}^{2n+1}((\partial E \setminus \partial^* E) \cap \Omega) = 0.$$

Theorem A.3. *Let $(E_j)_{j\in\mathbb{N}}$ be a sequence of (Λ, r)-minimizers of H-perimeter in an open set $\Omega \subset \mathbb{H}^n$, $\Lambda r \leq 1$. Then there exists a (Λ, r)-minimizer E of H-perimeter in Ω and a subsequence $(E_{j_k})_{k\in\mathbb{N}}$ such that*

$$E_{j_k} \longrightarrow E \quad \text{in } L^1_{\mathrm{loc}}(\Omega) \qquad \text{and} \qquad \nu_{E_{j_k}}\mu_{E_{j_k}} \xrightarrow{\;*\;} \nu_E \mu_E$$

as $k \to \infty$. Moreover, the measure-theoretic boundaries ∂E_{j_k} converge to ∂E in the sense of Kuratowski, i.e.,

(i) *if $p_{j_k} \in \partial E_j \cap \Omega$ and $p_{j_k} \to p \in \Omega$, then $p \in \partial E$;*

(ii) *if $p \in \partial E \cap \Omega$, then there exists a sequence $(p_{j_k})_{k\in\mathbb{N}}$ such that $p_{j_k} \in \partial E_{j_k} \cap \Omega$ and $p_{j_k} \to p$.*

For a proof in the case of the perimeter in $\mathbb{R}^n$, see [Maggi 2012, Chapter 21].

Appendix B

We define a Borel unit normal ν_R to an $\mathscr{S}^{2n+1}$-rectifiable set $R \subset \mathbb{H}^n$ and we show that the definition is well posed $\mathscr{S}^{2n+1}$-a.e., up to the sign. The normal ν_S to an H-regular hypersurface $S \subset \mathbb{H}^n$ is defined in (1-6).

Definition B.1. Let $R \subset \mathbb{H}^n$ be an $\mathscr{S}^{2n+1}$-rectifiable set such that

$$\mathscr{S}^{2n+1}\Big(R \setminus \bigcup_{j\in\mathbb{N}} S_j\Big) = 0 \tag{B-92}$$

for a sequence of H-regular hypersurfaces $(S_j)_{j\in\mathbb{N}}$ in $\mathbb{H}^n$. For any $p \in R \cap \bigcup_{j\in\mathbb{N}} S_j$, we define

$$\nu_R(p) = \nu_{S_{\bar{j}}}(p),$$

where $\bar{j}$ is the unique integer such that $p \in S_{\bar{j}} \setminus \bigcup_{j<\bar{j}} S_j$.

We show that Definition B.1 is well posed, up to a sign, for $\mathscr{S}^{2n+1}$-a.e. p. Namely, let $(S_j^1)_{j\in\mathbb{N}}$ and $(S_j^2)_{j\in\mathbb{N}}$ be two sequences of H-regular hypersurfaces in $\mathbb{H}^n$ for which (B-92) holds and denote by ν_R^1 and ν_R^2, respectively, the associated normals to R according to Definition B.1. We show that $\nu_R^1 = \nu_R^2$ $\mathscr{S}^{2n+1}$-a.e. on R, up to the sign.

Let $A \subset R$ be the set of points such that either $\nu_R^1(p)$ is not defined, or $\nu_R^2(p)$ is not defined, or they are both defined and $\nu_R^1(p) \neq \pm\nu_R^2(p)$. It is enough to show that $\mathscr{S}^{2n+1}(A) = 0$. This is a consequence of the following lemma:

Lemma B.2. *Let S_1, S_2 be two H-regular hypersurfaces in $\mathbb{H}^n$ and let*

$$A = \{p \in S_1 \cap S_2 : \nu_{S_1}(p) \neq \pm\nu_{S_2}(p)\}.$$

Then, the Hausdorff dimension of A in the Carnot–Carathéodory metric is at most $2n$, $\dim_{CC}(A) \le 2n$, and, in particular, $\mathscr{S}^{2n+1}(A) = 0$.

Proof. The blow-up of S_i, $i = 1, 2$, at a point $p \in A$ is a vertical hyperplane $\Pi_i \times \mathbb{R} \subset \mathbb{R}^{2n} \times \mathbb{R} \equiv \mathbb{H}^n$ — see, e.g., [Franchi et al. 2001] — where:

(i) By blow-up of S_i at p, we mean the limit

$$\lim_{\lambda\to\infty} \lambda(p^{-1} * S_i)$$

in the Gromov–Hausdorff sense. Recall that, for $E \subset \mathbb{H}^n$, we define $\lambda E = \{(\lambda z, \lambda^2 t) \in \mathbb{H}^n : (z, t) \in E\}$).

(ii) For $i = 1, 2$, $\Pi_i \subset \mathbb{R}^{2n}$ is the normal hyperplane to $\nu_{S_i}(p) \in H_p \equiv \mathbb{R}^{2n}$.

It follows that the blow-up of A at p is contained in the blow-up of $S_1 \cap S_2$ at p, i.e., in $(\Pi_1 \cap \Pi_2) \times \mathbb{R}$. Since $\nu_{S_1}(p) \neq \pm\nu_{S_2}(p)$, $\Pi_1 \cap \Pi_2$ is a $(2n-2)$-dimensional plane in $\mathbb{R}^{2n}$, and we conclude thanks to the following lemma. $\square$

Lemma B.3. *Let $k = 0, 1, \dots, 2n$ and $A \subset \mathbb{H}^n$ be such that, for any $p \in A$, the blow-up of A at p is contained in $\Pi_p \times \mathbb{R}$ for some plane $\Pi_p \subset \mathbb{R}^{2n}$ of dimension k. Then we have* $\dim_{CC}(A) \leq k + 2$.

Proof. We claim that, for any $\eta > 0$, we have

$$\mathscr{S}^{k+2+\eta}(A) = 0. \tag{B-93}$$

Let $\varepsilon \in \left(0, \frac{1}{2}\right)$ be such that $C\varepsilon^\eta \leq \frac{1}{2}$, where $C = C(n)$ is a constant that will be fixed later in the proof. By the definition of blow-up, for any $p \in A$ there exists $r_p > 0$ such that, for all $r \in (0, r_p)$, we have

$$(p^{-1} * A) \cap U_r \subset (\Pi_p)_{\varepsilon r} \times \mathbb{R},$$

where $(\Pi_p)_{\varepsilon r}$ denotes the (εr)-neighbourhood of Π_p in $\mathbb{R}^{2n}$. For any $j \in \mathbb{N}$, set

$$A_j = \{p \in A \cap B_j : r_p > 1/j\}.$$

To prove (B-93), it is enough to prove that

$$\mathscr{S}^{k+2+\eta}(A_j) = 0$$

for any fixed $j \geq 1$. This, in turn, will follow if we show that, for any fixed $\delta \in (0, 1/(2j))$, one has

$$\inf\left\{\sum_{i\in\mathbb{N}} r_i^{k+2+\eta} : A_j \subset \bigcup_{i\in\mathbb{N}} U_{r_i}(p_i),\ r_i < 2\varepsilon\delta\right\} \leq \frac{1}{2} \inf\left\{\sum_{i\in\mathbb{N}} r_i^{k+2+\eta} : A_j \subset \bigcup_{i\in\mathbb{N}} U_{r_i}(p_i),\ r_i < \delta\right\}. \tag{B-94}$$

Let $(U_{r_i}(p_i))_{i\in\mathbb{N}}$ be a covering of A_j with balls of radius smaller than δ. There exist points $\bar{p}_i \in A_j$ such that $(U_{2r_i}(\bar{p}_i))_{i\in\mathbb{N}}$ is a covering of A_j with balls of radius smaller than $2\delta < 1/j$. By definition of A_j, we have

$$(\bar{p}_i^{-1} * A_j) \cap U_{2r_i} \subset ((\Pi_{\bar{p}_i})_{\varepsilon r_i} \times \mathbb{R}) \cap U_{2r_i}.$$

The set $((\Pi_{\bar{p}_i})_{\varepsilon r_i} \times \mathbb{R}) \cap U_{2r_i}$ can be covered by a family of balls $(U_{\varepsilon r_i}(p_h^i))_{h\in H_i}$ of radius $\varepsilon r_i < 2\varepsilon\delta$ in such a way that the cardinality of H_i is bounded by $C\varepsilon^{-k-2}$, where the constant C depends only on n and not on ε. In particular, the family of balls $(U_{\varepsilon r_i}(\bar{p}_i * p_h^i))_{i\in\mathbb{N}, h\in H_i}$ is a covering of A_j and

$$\begin{aligned}\sum_{i\in\mathbb{N}} \sum_{h\in H_i} (\text{radius } U_{\varepsilon r_i}(\bar{p}_i * p_h^i))^{k+2+\eta} &= \sum_{i\in\mathbb{N}} \sum_{h\in H_i} (\varepsilon r_i)^{k+2+\eta} \leq C\varepsilon^{-k-2} \sum_{i\in\mathbb{N}} (\varepsilon r_i)^{k+2+\eta} \\ &= C\varepsilon^\eta \sum_{i\in\mathbb{N}} r_i^{k+2+\eta} \leq \frac{1}{2} \sum_i r_i^{k+2+\eta}.\end{aligned}$$

This proves (B-94) and concludes the proof. $\square$

References

[Almgren 1968] F. J. Almgren, Jr., "Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure", *Ann. of Math.* (2) **87** (1968), 321–391. MR 37 #837 Zbl 0162.24703

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Clarendon/Oxford University Press, New York, 2000. MR 2003a:49002 Zbl 0957.49001

[Ambrosio et al. 2006] L. Ambrosio, F. Serra Cassano, and D. Vittone, "Intrinsic regular hypersurfaces in Heisenberg groups", *J. Geom. Anal.* **16**:2 (2006), 187–232. MR 2007g:49072 Zbl 1085.49045

[Burago and Zalgaller 1988] Y. D. Burago and V. A. Zalgaller, *Geometric inequalities*, Grundlehren der Math. Wissenschaften **285**, Springer, Berlin, 1988. MR 89b:52020 Zbl 0633.53002

[Federer 1969] H. Federer, *Geometric measure theory*, Grundlehren der Math. Wissenschaften **153**, Springer, New York, 1969. MR 41 #1976 Zbl 0176.00801

[Franchi et al. 1996] B. Franchi, R. Serapioni, and F. Serra Cassano, "Meyers–Serrin type theorems and relaxation of variational integrals depending on vector fields", *Houston J. Math.* **22**:4 (1996), 859–890. MR 98c:49037 Zbl 0876.49014

[Franchi et al. 2001] B. Franchi, R. Serapioni, and F. Serra Cassano, "Rectifiability and perimeter in the Heisenberg group", *Math. Ann.* **321**:3 (2001), 479–531. MR 2003g:49062 Zbl 1057.49032

[Franchi et al. 2007] B. Franchi, R. Serapioni, and F. Serra Cassano, "Regular submanifolds, graphs and area formula in Heisenberg groups", *Adv. Math.* **211**:1 (2007), 152–203. MR 2008h:49030 Zbl 1125.28002

[Garofalo and Nhieu 1996] N. Garofalo and D.-M. Nhieu, "Isoperimetric and Sobolev inequalities for Carnot–Carathéodory spaces and the existence of minimal surfaces", *Comm. Pure Appl. Math.* **49**:10 (1996), 1081–1144. MR 97i:58032 Zbl 0880.35032

[Hajłasz and Koskela 2000] P. Hajłasz and P. Koskela, *Sobolev met Poincaré*, vol. 145, Mem. Amer. Math. Soc. **688**, American Math. Society, Providence, RI, 2000. MR 2000j:46063 Zbl 0954.46022

[Heinonen 2001] J. Heinonen, *Lectures on analysis on metric spaces*, Springer, New York, 2001. MR 2002c:30028 Zbl 0985.46008

[Kirchheim and Serra Cassano 2004] B. Kirchheim and F. Serra Cassano, "Rectifiability and parameterization of intrinsic regular surfaces in the Heisenberg group", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **3**:4 (2004), 871–896. MR 2005m:28013 Zbl 1170.28300

[Maggi 2012] F. Maggi, *Sets of finite perimeter and geometric variational problems: an introduction to geometric measure theory*, Cambridge Studies in Advanced Mathematics **135**, Cambridge University Press, Cambridge, 2012. MR 2976521 Zbl 1255.49074

[Magnani 2004] V. Magnani, "Note on coarea formulae in the Heisenberg group", *Publ. Mat.* **48**:2 (2004), 409–422. MR 2005d:28014 Zbl 1062.22020

[Magnani 2008] V. Magnani, "Non-horizontal submanifolds and coarea formula", *J. Anal. Math.* **106** (2008), 95–127. MR 2009k:53067 Zbl 1158.22006

[Magnani 2014] V. Magnani, "A new differentiation, shape of the unit ball and perimeter measure", preprint, 2014. arXiv 1408.4726

[Mattila et al. 2010] P. Mattila, R. Serapioni, and F. Serra Cassano, "Characterizations of intrinsic rectifiability in Heisenberg groups", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **9**:4 (2010), 687–723. MR 2012c:28003 Zbl 1229.28004

[Monti 2014] R. Monti, "Lipschitz approximation of $\mathbb{H}$-perimeter minimizing boundaries", *Calc. Var. Partial Differential Equations* **50**:1–2 (2014), 171–198. MR 3194680 Zbl 1290.49088

[Monti 2015] R. Monti, "Minimal surfaces and harmonic functions in the Heisenberg group", *Nonlinear Analysis* (online publication April 2015).

[Monti and Morbidelli 2005] R. Monti and D. Morbidelli, "Regular domains in homogeneous groups", *Trans. Amer. Math. Soc.* **357**:8 (2005), 2975–3011. MR 2006b:43016 Zbl 1067.43003

[Monti and Vittone 2012] R. Monti and D. Vittone, "Sets with finite $\mathbb{H}$-perimeter and controlled normal", *Math. Z.* **270**:1–2 (2012), 351–367. MR 2875838 Zbl 1257.49047

[Schoen and Simon 1982] R. Schoen and L. Simon, "A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals", *Indiana Univ. Math. J.* **31**:3 (1982), 415–434. MR 84j:49039 Zbl 0516.49026

[Serra Cassano and Vittone 2014] F. Serra Cassano and D. Vittone, "Graphs of bounded variation, existence and local boundedness of non-parametric minimal surfaces in Heisenberg groups", *Adv. Calc. Var.* **7**:4 (2014), 409–492. MR 3276118 Zbl 1304.49084

[Vittone 2012] D. Vittone, "Lipschitz surfaces, perimeter and trace theorems for BV functions in Carnot–Carathéodory spaces", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **11**:4 (2012), 939–998. MR 3060706 Zbl 1270.53068

Received 7 Oct 2014. Accepted 11 May 2015.

ROBERTO MONTI: monti@math.unipd.it
Dipartimento di Matematica, Universitá di Padova, via Trieste 63, I-35121 Padova, Italy

DAVIDE VITTONE: vittone@math.unipd.it
Dipartimento di Matematica, Universitá di Padova, via Trieste 63, I-35121 Padova, Italy

and

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

ANALYSIS AND PDE
Vol. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1455

IMPROVEMENT OF THE ENERGY METHOD FOR STRONGLY NONRESONANT DISPERSIVE EQUATIONS AND APPLICATIONS

LUC MOLINET AND STÉPHANE VENTO

We propose a new approach to prove the local well-posedness of the Cauchy problem associated with strongly nonresonant dispersive equations. As an example, we obtain unconditional well-posedness of the Cauchy problem in the energy space for a large class of one-dimensional dispersive equations with a dispersion that is greater than the one of the Benjamin–Ono equation. At the level of dispersion of the Benjamin–Ono, we also prove the well-posedness in the energy space but without unconditional uniqueness. Since we do not use a gauge transform, this enables us in all cases to prove strong convergence results in the energy space for solutions of viscous versions of these equations towards the purely dispersive solutions. Finally, it is worth noting that our method of proof works on the torus as well as on the real line.

1. Introduction

The Cauchy problem associated with dispersive equations with derivative nonlinearity has been extensively studied since the eighties. The first results were obtained by using energy methods that did not make use of the dispersive effects (see for instance [Kato 1983; Abdelouhab et al. 1989]). These methods were restricted to regular initial data ($s > d/2$, where $d \geq 1$ is the spatial dimension) and only ensured the continuity of the solution map. At the end of the eighties, Kenig, Ponce and Vega proved new dispersive estimates that enable them to lower the regularity requirement on the initial data (see for instance [Kenig et al. 1991; 1993; Ponce 1991]). They even obtained local well-posedness (LWP) for a large class of dispersive equations by a fixed point argument in a suitable Banach space related to linear dispersive estimates. Then, Bourgain [1993a; 1993b] introduced the now so-called Bourgain spaces, where one can solve by a fixed point argument a wide class of dispersive equations with very rough initial data. It is worth noting that, since the nonlinearity of these equations is in general algebraic, the fixed point argument ensures the real analyticity of the solution map. Molinet, Saut and Tzvetkov [Molinet et al. 2001] noticed that a large class of "weakly" dispersive equations, including in particular the Benjamin–Ono equation, cannot be solved by a fixed point argument for initial data in any Sobolev spaces H^s. This obstruction is due to bad interactions between high frequencies and very low frequencies. Since then, roughly speaking, two approaches have been developed to lower the regularity requirement for such equations. The first one is the so-called gauge method. This consists in introducing a nonlinear gauge transform of the solution that solved an equation with fewer bad interactions than the original one. This method proved to be very

MSC2010: 35E15, 35Q53, 35A02.
Keywords: Benjamin–Ono equation, intermediate long wave equation, dispersion generalized Benjamin–Ono equation, well-posedness, unconditional uniqueness.

efficient for obtaining the lowest regularity index for solving canonical equations (see [Tao 2004; Ionescu and Kenig 2007; Burq and Planchon 2008; Molinet and Pilod 2012] for the BO equation and [Herr et al. 2010] for the dispersive generalized BO equation) but has the disadvantage of behaving very badly with respect to perturbation of the equation. The second one consists in improving the dispersive estimates by localizing it in space-frequency-depending time intervals and then mixing it with classical energy estimates. This type of method was first introduced by Koch and Tzvetkov [2003] (see also [Kenig and Koenig 2003] for some improvements) in the framework of Strichartz's spaces and then by Koch and Tataru [2007] (see also [Ionescu et al. 2008]) in the framework of Bourgain's spaces. It is less efficient for getting the best regularity index but it is surely more flexible with respect to perturbation of the equation.

In this paper we propose a new approach to derive local and global well-posedness results for dispersive equations that do not exhibit too-strong resonances. This approach combines classical energy estimates with Bourgain-type estimates on a time interval that does not depend on the space frequency. Here, we will apply this method to prove unconditional local well-posedness results on both $\mathbb{R}$ and $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ without the use of a gauge transform for a large class of one-dimensional quadratic dispersive equations with a dispersion between those of the Benjamin–Ono equation and the KdV equation. This class contains, in particular, the equations with pure power dispersion that read

$$u_t + \partial_x D_x^\alpha u + u u_x = 0 \tag{1-1}$$

with $\alpha \in [1, 2]$.

The principle of the method is particularly simple in the regular case $s > \frac{1}{2}$. We start with the classical space-frequency-localized energy estimate

$$\|P_N u\|_{L_T^\infty H^s}^2 \lesssim \|P_N u_0\|_{H^s}^2 + \sup_{t \in]0,T[} \langle N\rangle^{2s} \left| \int_0^t \int_{\mathbb{R}} \partial_x P_N(u^2) P_N u \right|, \tag{1-2}$$

obtained by projecting the equation on frequencies of order N and taking the inner product with $J_x^s u$. Note that the second term in the right-hand side of (1-2) is easily controlled (after summing in N) by $\|u\|_{L_T^\infty H^s}^3$ for $s > \frac{3}{2}$. This is the main point in the standard energy method that leads to LWP in H^s, $s > \frac{3}{2}$. In order to take into account the dispersive effects of the equation, we will decompose the three factors in the integral term into dyadic pieces for the modulation variables and use the Bourgain spaces $X^{s,b}$ in a nonconventional way. Actually, it is known that standard bilinear estimates in $X^{s,b}$-spaces with $b = \frac{1}{2}+$ fail for (1-1) for any $s \in \mathbb{R}$ as soon as $\alpha < 2$. On the other hand, as noticed in [Zhou 1997], it is easy to deduce from the equation that a solution $u \in L^\infty(0, T; H^s)$ to (1-3) has to belong to the space $X_T^{s-1,1}$. This means that, if we accept the loss of a few spatial derivatives on the solution, then we may gain some regularity in the modulation variable. This is particularly profitable when the equation enjoys a strong nonresonance relation such as (2-6). Actually, this formally allows us to estimate the second term in (1-2) at the desired level. However, this term involves a multiplication by $1_{]0,t[}$ and it is well known that such multiplication is not bounded in $X^{s-1,1}$. To overcome this difficulty we decompose this function into two parts: a high-frequency part that will be very small in L_T^1 and a low-frequency part that will have good properties with respect to multiplication with high-modulation functions in $X^{s-1,1}$. This decomposition will depend on the space-frequency-localization of the three functions that appear in the trilinear term.

1A. ***Presentation of the results.*** In this paper we consider the dispersive equation

$$u_t + L_{\alpha+1}u + \tfrac{1}{2}\partial_x(u^2) = 0 \tag{1-3}$$

associated with the initial condition

$$u(0,\cdot) = u_0, \tag{1-4}$$

where $x \in \mathbb{R}$ or $\mathbb{T}$, $u = u(t,x)$ and $u_0 = u_0(x)$ are real-valued functions, $\alpha > 0$ is a real number and the linear operator $L_{\alpha+1}$ satisfies the following hypothesis:

Hypothesis 1. *$L_{\alpha+1}$ is the Fourier multiplier operator by $i p_{\alpha+1}$, where $p_{\alpha+1}$ is a real-valued odd function satisfying, for some $\lambda_0 > 0$,*

(1) *For any $|\xi| \gg 1$ and $0 < \lambda \le \lambda_0$,*

$$\lambda^{\alpha+1}|p_{\alpha+1}(\lambda^{-1}\xi)| \lesssim |\xi|^{\alpha+1}. \tag{1-5}$$

(2) *For any $(\xi_1,\xi_2) \in \mathbb{R}^2$ with $|\xi_1| \gg 1$ and any $0 < \lambda \le \lambda_0$,*

$$\lambda^{\alpha+1}|\Omega(\lambda^{-1}\xi_1, \lambda^{-1}\xi_2)| \sim |\xi|_{\min}|\xi|_{\max}^{\alpha}, \tag{1-6}$$

where

$$\begin{aligned} \Omega(\xi_1,\xi_2) &:= p_{\alpha+1}(\xi_1+\xi_2) - p_{\alpha+1}(\xi_1) - p_{\alpha+1}(\xi_2), \\ |\xi|_{\min} &:= \min(|\xi_1|, |\xi_2|, |\xi_1+\xi_2|) \\ \textit{and} \quad |\xi|_{\max} &:= \max(|\xi_1|, |\xi_2|, |\xi_1+\xi_2|). \end{aligned}$$

Remark 1.1. We will see in Lemma 2.1 below that, for $\alpha > 0$, a very simple criterion on p ensures (1-6). With this criterion in hand, it is not too hard to check that the following linear operators satisfy Hypothesis 1:

(1) The purely dispersive operators $L := \partial_x D_x^\alpha$ with $\alpha > 0$.

(2) The linear intermediate long wave operator $L := \partial_x D_x \coth D_x$. Note that here $\alpha = 1$.

(3) Some perturbations of the Benjamin–Ono equation, such as the Smith operator [1972], $L := \partial_x(D_x^2+1)^{1/2}$. Here again $\alpha = 1$.

Before stating our main result, let us define what we mean by unconditional well-posedness.

Definition 1.2. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$, $T > 0$ and $s \ge 0$. We will say that $u \in L^\infty(0,T;H^s(\mathbb{K}))$ is a solution to (1-3) associated with the initial datum $u_0 \in H^s(\mathbb{K})$ if u satisfies (1-3)–(1-4) in the distributional sense, i.e., for any test function $\phi \in C_c^\infty(]-T,T[\times\mathbb{K})$,

$$\int_0^\infty \int_{\mathbb{K}} \left[(\phi_t + L_{\alpha+1}\phi)u + \tfrac{1}{2}\phi_x u^2\right] dx\,dt + \int_{\mathbb{K}} \phi(0,\cdot)u_0\,dx = 0 \tag{1-7}$$

Remark 1.3. For $u \in L^\infty(0,T;H^s(\mathbb{K}))$, with $s \ge 0$, u^2 is well defined and is in $L^\infty(0,T;H^{s-(1/2+)}(\mathbb{K}))$. Moreover, (1-5) forces

$$L_{\alpha+1}u \in L^\infty(0,T;H^{s-\alpha-1}(\mathbb{K})).$$

Therefore, $u_t \in L^\infty(0,T;H^{s-\alpha-1}(\mathbb{K}))$ and (1-7) ensures that (1-3) is satisfied in $L^\infty(0,T;H^{s-\alpha-1}(\mathbb{K}))$. In particular, $u \in C([0,T];H^{s-\alpha-1}(\mathbb{K}))$ and (1-7) forces the initial condition $u(0) = u_0$. Note that this actually implies that $u \in C([0,T]; H^\theta(\mathbb{K}))$ for any $\theta < s$. Finally, we note that this ensures that u satisfies the Duhamel formula associated with (1-3).

Definition 1.4. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$ and $s \in \mathbb{R}$. We will say that the Cauchy problem associated with (1-3) is unconditionally locally well-posed in $H^s(\mathbb{K})$ if, for any initial data $u_0 \in H^s(\mathbb{K})$, there exists $T = T(\|u_0\|_{H^s}) > 0$ and a solution $u \in C([0,T]; H^s(\mathbb{K}))$ to (1-3) emanating from u_0. Moreover, u is the unique solution to (1-3) associated with u_0 that belongs to $L^\infty(]0,T[; H^s(\mathbb{K}))$. Finally, for any $R > 0$, the solution map $u_0 \mapsto u$ is continuous from the ball of $H^s(\mathbb{K})$ with radius R centered at the origin into $C([0,T(R)]; H^s(\mathbb{K}))$.

Theorem 1.5. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$, $L_{\alpha+1}$ satisfy Hypothesis 1 with $1 \le \alpha \le 2$ and let $s \ge 1 - \frac{\alpha}{2}$ with $(s,\alpha) \ne \left(\frac{1}{2}, 1\right)$. Then the Cauchy problem associated with (1-3) is unconditionally locally well-posed in $H^s(\mathbb{K})$ with a maximal time of existence $T \gtrsim (1 + \|u_0\|_{H^{1-\alpha/2}})^{-2(\alpha+1)/(2\alpha-1)}$.*

Remark 1.6. In the regular case (Cauchy problem in H^s with $s > \frac{1}{2}$), we actually need (1-6) only for $|\xi_1| \wedge |\xi_2| \gg 1$.

Remark 1.7. Our method also works in the case $\alpha > 2$. In this case we get the unconditional well-posedness in $H^s(\mathbb{K})$ for $s \ge 0$.

Remark 1.8. For $L_{\alpha+1} := \partial_x^3$, we recover the unconditional LWP results for the KdV equation in $L^2(\mathbb{R})$ and $L^2(\mathbb{T})$ obtained in [Zhou 1997; Babin et al. 2011], respectively.

For $L_{\alpha+1}$ with $\alpha \in]1,2[$ our results on unconditional uniqueness are, to our knowledge, new for both the real line case and the periodic case.

In the limit case $(s,\alpha) = \left(\frac{1}{2}, 1\right)$ we do not succeed in proving the unconditional uniqueness result. However, our method of proof enables us to prove the well-posedness without using a gauge transform. We think that this result is also of interest since $H^{1/2}$ is the energy space when $\alpha = 1$. It is worth noticing that, as far as we know, the available results without gauge transformation on the local well-posedness of the Benjamin–Ono equation in Sobolev spaces $H^s(\mathbb{R})$ were restricted to $s \ge 1$ (see [Guo et al. 2011]). Also, the well-posedness in the energy space $H^{1/2}$ seems to be new for the intermediate long waves equation, at least in the periodic setting.

Theorem 1.9. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$ and assume $L_{\alpha+1}$ satisfies Hypothesis 1 with $\alpha = 1$. Then the Cauchy problem associated with (1-3) is locally well-posed in $H^{1/2}(\mathbb{K})$ with a maximal time of existence $T \gtrsim (1 + \|u_0\|_{H^{1/2}})^{-4}$.*

Let us assume now that the symbol $p_{\alpha+1}$ satisfies, moreover,

$$|p_{\alpha+1}(\xi)| \lesssim |\xi| \quad \text{for } |\xi| \le 1 \qquad \text{and} \qquad |p_{\alpha+1}(\xi)| \sim |\xi|^{\alpha+1} \quad \text{for } |\xi| \ge 1. \tag{1-8}$$

Then it is not too hard to check that (1-3) enjoys the conservation laws

$$\frac{d}{dt}\int_{\mathbb{K}} u^2\,dx = 0 \quad \text{and} \quad \frac{d}{dt}\int_{\mathbb{K}} \left(|\Lambda^{\alpha/2}u|^2 + \tfrac{1}{3}u^3\right)dx = 0,$$

where $\Lambda^{\alpha/2}$ is the space Fourier multiplier defined by

$$\widehat{\Lambda^{\alpha/2} v}(\xi) = \left| \frac{p_{\alpha+1}(\xi)}{\xi} \right|^{\frac{1}{2}} \hat{v}(\xi).$$

Combined with the embedding $H^{\alpha/2} \hookrightarrow L^3$, we get an a priori bound of the $H^{\alpha/2}$-norm of the solution. We may then iterate Theorems 1.5 and 1.9 to obtain the following corollary:

Corollary 1.10. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$ and assume $L_{\alpha+1}$ satisfies Hypothesis 1 and (1-8). Then the Cauchy problem associated with (1-3) is unconditionally globally well-posed in $H^{\alpha/2}(\mathbb{K})$ for $1 < \alpha \leq 2$, and globally well-posed in $H^{1/2}(\mathbb{T})$ for $\alpha = 1$.*

Remark 1.11. The linear operators given in Remark 1.1 also satisfy assumption (1-8).

Remark 1.12. If one considers LWP and not unconditional LWP, then the best-known results for (1-1) with $1 < \alpha < 2$ have been obtained in [Herr et al. 2010], where the global well-posedness in $L^2(\mathbb{R})$ is proved by using a paradifferential gauge transformation. As far as we know, the best available results without gauge transformation are obtained in [Guo 2012], where the LWP in $H^s(\mathbb{R})$ with $s > 2-\alpha$ is proven. This leads to a global well-posedness result only for $\alpha > \frac{4}{3}$. Therefore, even for (1-1), our results improve the local and global available well-posedness results without the use of gauge transformation on the real line. To the best of our knowledge, they are new on the one-dimensional torus, where we are not aware of any global well-posedness result.

It is well known that, taking into account some damping or dissipative effects, dissipative versions of (1-3) can be derived (see for instance [Ott and Sudan 1970; Edwin and Roberts 1986]). One quite direct application of the fact that we do not need a gauge transform to solve (1-3) is that we can easily treat the dissipative limit of dissipative versions of (1-3). Such a dissipative limit was, for example, studied for the Benjamin–Ono equation on the real line in [Guo et al. 2011; Molinet 2013].

Let us introduce the following dissipative version of (1-3):

$$u_t + L_{\alpha+1} u + \varepsilon A_\beta u + u u_x = 0, \tag{1-9}$$

where $\varepsilon > 0$ is a small parameter, $\beta \geq 0$ and A_β is a linear operator satisfying the following hypothesis:

Hypothesis 2. *We assume that A_β is the Fourier multiplier operator by q_β, where q_β is a real-valued, even function, bounded on bounded intervals, satisfying: for all $0 < \lambda \ll 1$ and $\xi \gg 1$,*

$$\lambda^\beta q_\beta(\lambda^{-1}\xi) \sim |\xi|^\beta.$$

Remark 1.13. Both the homogeneous operators D_x^β and the nonhomogeneous operators J_x^β satisfy Hypothesis 2.

Theorem 1.14. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{T}$, $1 \leq \alpha \leq 2$, $0 \leq \beta \leq 1+\alpha$ and $s \geq 1 - \frac{\alpha}{2}$.*

(1) *Then the Cauchy problem associated with (1-9) is locally well-posed in $H^s(\mathbb{K})$.*

(2) *For* $u_0 \in H^s(\mathbb{K})$, *let* u *be the solution to* (1-3) *emanating from* u_0 *and let the maximal time of existence of* u *in* H^s *be* $T^* \gtrsim (1 + \|u_0\|_{H^{1-\alpha/2}})^{-2(\alpha+1)/(2\alpha-1)}$ (*note that* T^* *may be infinite*). *Then the maximal time of existence* T_ε *of the solution* u_ε *to* (1-9) *emanating from* u_0 *satisfies* $\liminf_{\varepsilon\to 0} T_\varepsilon \geq T^*$. *Moreover, for any* $0 < T_0 < T^*$, $u_\varepsilon \to u$ *in* $C([0, T_0]; H^s)$ *as* $\varepsilon \to 0$.

Remark 1.15. The constraint $\beta \leq 1+\alpha$ is clearly an artifact of the method we used. We think that it could be dropped by replacing, in some estimates, the dispersive Bourgain spaces by dispersive–dissipative Bourgain spaces that were first introduced in [Molinet and Ribaud 2002]. But, since the dissipative operators involved in wave motions are commonly of order less or equal to 2, we do not pursue this issue.

The rest of the paper is organized as follows: In Section 2, we introduce the notations, define the function spaces and state some preliminary lemmas. In Section 3, we develop our method in the simplest case, $s > \frac{1}{2}$, while the nonregular case is treated in Section 4. Section 5 is devoted to the proof of Theorem 1.14. We conclude the paper with an Appendix explaining how to deal with the special case $(s,\alpha) = \left(\frac{1}{2}, 1\right)$.

2. Notations, function spaces and preliminary lemmas

2A. ***Notation.*** For any positive numbers a and b, the notation $a \lesssim b$ means that there exists a positive constant c such that $a \leq cb$. We also write $a \sim b$ when $a \lesssim b$ and $b \lesssim a$. Moreover, if $\alpha \in \mathbb{R}$, then α_+ and α_- will denote a number slightly greater and less than α, respectively.

For $u = u(x,t) \in \mathcal{S}(\mathbb{R}^2)$, $\mathcal{F}u = \hat{u}$ will denote its space-time Fourier transform, whereas $\mathcal{F}_x u = (u)^{\wedge_x}$ and $\mathcal{F}_t u = (u)^{\wedge_t}$ will denote its Fourier transform in space and in time, respectively. For $s \in \mathbb{R}$, we define the Bessel and Riesz potentials of order $-s$, J_x^s and D_x^s, by

$$J_x^s u = \mathcal{F}_x^{-1}\big((1+|\xi|^2)^{\frac{s}{2}}\mathcal{F}_x u\big) \quad \text{and} \quad D_x^s u = \mathcal{F}_x^{-1}(|\xi|^s \mathcal{F}_x u).$$

Throughout the paper, we fix a smooth cutoff function η such that

$$\eta \in C_0^\infty(\mathbb{R}), \quad 0 \leq \eta \leq 1, \quad \eta_{|[-1,1]} = 1 \quad \text{and} \quad \operatorname{supp}(\eta) \subset [-2,2].$$

We set $\phi(\xi) := \eta(\xi) - \eta(2\xi)$. For $l \in \mathbb{Z}$, we define

$$\phi_{2^l}(\xi) := \phi(2^{-l}\xi),$$

and, for $l \in \mathbb{N}^*$,

$$\psi_{2^l}(\xi,\tau) = \phi_{2^l}(\tau - p_{\alpha+1}(\xi)),$$

where $i p_{\alpha+1}$ is the Fourier symbol of $L_{\alpha+1}$. By convention, we also denote

$$\psi_1(\xi,\tau) := \eta\big(2(\tau - p_{\alpha+1}(\xi))\big).$$

Any summations over capitalized variables such as N, L, K or M are presumed to be dyadic. Unless stated otherwise, we work with homogeneous dyadic decomposition for the space-frequency variables and nonhomogeneous decompositions for modulation variables, i.e., these variables range over numbers of the form $\{2^k : k \in \mathbb{Z}\}$ and $\{2^k : k \in \mathbb{N}\}$, respectively. Then we have that

$$\sum_{N>0} \phi_N(\xi) = 1 \quad \text{for all } \xi \in \mathbb{R}^* \qquad \text{and} \qquad \operatorname{supp}(\phi_N) \subset \left\{\tfrac{1}{2}N \le |\xi| \le 2N\right\} \quad \text{for } N \in \{2^k : k \in \mathbb{Z}\},$$

and

$$\sum_{L\ge 1} \psi_L(\xi,\tau) = 1 \quad \text{for all } (\xi,\tau) \in \mathbb{R}^2$$

Let us define the Littlewood–Paley multipliers by

$$P_N u = \mathcal{F}_x^{-1}(\phi_N \mathcal{F}_x u), \quad Q_L u = \mathcal{F}^{-1}(\psi_L \mathcal{F} u),$$

$P_{\ge N} := \sum_{K\ge N} P_K$, $P_{\le N} := \sum_{K\le N} P_K$, $Q_{\ge L} := \sum_{K\ge L} Q_K$ and $Q_{\le L} := \sum_{K\le L} Q_K$. For brevity we also write $u_N = P_N u$, $u_{\le N} = P_{\le N} u$, etc.

Let χ be a (possibly complex-valued) bounded measurable function on $\mathbb{R}^2$ and define the pseudoproduct operator $\Pi = \Pi_\chi$ on $\mathcal{S}(\mathbb{R})^2$ by

$$\mathcal{F}(\Pi(f,g))(\xi) = \int_{\mathbb{R}} \hat{f}(\xi_1)\hat{g}(\xi-\xi_1)\chi(\xi,\xi_1)\,d\xi_1.$$

Throughout the paper, we write $\Pi = \Pi_\chi$, where χ may be different at each occurrence of Π. This bilinear operator behaves like a product in the sense that it satisfies the following properties:

$$\Pi(f,g) = fg \quad \text{if } \chi \equiv 1,$$
$$\int_{\mathbb{R}} \Pi_\chi(f,g)h = \int_{\mathbb{R}} f\,\Pi_{\chi_1}(g,h) = \int_{\mathbb{R}} \Pi_{\chi_2}(f,h)g \tag{2-1}$$

with $\chi_1(\theta,\theta_1) = \bar{\chi}(\theta_1,\theta)$ and $\chi_2(\theta,\theta_1) = \bar{\chi}(\theta-\theta_1,\theta)$ for any real-valued functions f, g, $h \in \mathcal{S}(\mathbb{R})$. Moreover, we easily check from the Bernstein inequality that, if $f_i \in L^2(\mathbb{R})$ has a Fourier transform localized in an annulus $\{|\xi| \sim N_i\}$, $i = 1, 2, 3$, then

$$\left|\int_{\mathbb{R}} \Pi(f_1,f_2)f_3\right| \lesssim N_{\min}^{\frac{1}{2}} \prod_{i=1}^{3} \|f_i\|_{L^2}, \tag{2-2}$$

where the implicit constant only depends on $\|\chi\|_{L^\infty(\mathbb{R}^2)}$ and $N_{\min} = \min\{N_1,N_2,N_3\}$. With this notation in hand, we will be able to systematically estimate terms of the form

$$\int_{\mathbb{R}} P_N(u^2)\,\partial_x P_N u$$

to put the derivative on the lowest frequency factor.

2B. ***Function spaces.*** For $1 \le p \le \infty$, $L^p(\mathbb{R})$ is the usual Lebesgue space with the norm $\|\cdot\|_{L^p}$ and, for $s \in \mathbb{R}$, $H^s(\mathbb{R})$ is the usual Sobolev space with its usual norm,

$$\|\phi\|_{H^s} = \|J_x^s \phi\|_{L^2}.$$

If B is one of the spaces defined above, $1 \le p \le \infty$, we will define the space-time spaces $L^p_t B$ and $\widetilde{L}^p_t B$ equipped with the norms

$$\|f\|_{L^p_t B} = \left(\int_{\mathbb{R}} \|f(\cdot,t)\|_B^p \, dt \right)^{\frac{1}{p}},$$

with obvious modifications for $p = \infty$, and

$$\|f\|_{\widetilde{L}^p_t B} = \left(\sum_{N>0} \|P_N f\|^2_{L^p_t B} \right)^{\frac{1}{2}}.$$

For $s, b \in \mathbb{R}$, we introduce the Bourgain spaces $X^{s,b}$ related to the linear part of (1-3) as the completion of the Schwartz space $\mathcal{S}(\mathbb{R}^2)$ under the norm

$$\|v\|_{X^{s,b}} := \left(\int_{\mathbb{R}^2} \langle \tau - p_{\alpha+1}(\xi)\rangle^{2b} \langle \xi\rangle^{2s} |\hat{v}(\xi,\tau)|^2 \, d\xi \, d\tau \right)^{\frac{1}{2}}, \tag{2-3}$$

where $\langle x\rangle := 1 + |x|$ and $i p_{\alpha+1}$ is the Fourier symbol of $L_{\alpha+1}$. Recall that

$$\|v\|_{X^{s,b}} = \|U_\alpha(-t)v\|_{H^{s,b}_{t,x}},$$

where $U_\alpha(t) = \exp(tL_{\alpha+1})$ is the generator of the free evolution associated with (1-3).

Finally, we will use restriction-in-time versions of these spaces. Let $T > 0$ be a positive time and let Y be a normed space of space-time functions. The restriction space Y_T will be the space of functions $v : \mathbb{R} \times]0, T[\to \mathbb{R}$ satisfying

$$\|v\|_{Y_T} := \inf\{\|\tilde{v}\|_Y \mid \tilde{v} : \mathbb{R}\times\mathbb{R} \to \mathbb{R},\ \tilde{v}|_{\mathbb{R}\times]0,T[} = v\} < \infty.$$

2C. *Preliminary lemmas.*

Lemma 2.1. *Let $p : \mathbb{R} \to \mathbb{R}$ be an odd function belonging to $C^1(\mathbb{R}) \cap C^2(\mathbb{R}^*)$. Assume that there exist $\alpha > 0$ and $\xi_0 > 0$ such that, for all $\xi \ge \xi_0$,*

$$|p'(\xi)| \sim |\xi|^\alpha \quad \text{and} \quad |p''(\xi)| \sim |\xi|^{\alpha-1}. \tag{2-4}$$

Then the Fourier multiplier $L_{\alpha+1}$ by ip satisfies Hypothesis 1.

Proof. Let $0 < \lambda \le \xi_0^{-1}$ be a real number. First, by the mean value theorem, for $\xi \ge 1$,

$$|p(\lambda^{-1}\xi)| \lesssim |p(\xi_0)| + \lambda^{-(\alpha+1)}\xi^{\alpha+1} \lesssim \lambda^{-1}(\lambda\xi_0) \max_{\xi\in[0,\xi_0]} |p'(\xi)| + \xi^{\alpha+1}$$

and thus

$$\lambda^{\alpha+1}|p(\lambda^{-1}\xi)| \lesssim \lambda^\alpha \max_{\xi\in[0,\xi_0]} |p'(\xi)| + \xi^{\alpha+1} \lesssim \xi^{\alpha+1}$$

as soon as $\lambda \le \left(\max_{\xi\in[0,\xi_0]} |p'(\xi)|\right)^{-1/\alpha}$. This proves (1-5) for

$$\lambda_0 = \min\Big(\xi_0^{-1}, \big(\max_{\xi\in[0,\xi_0]} |p'(\xi)|\big)^{-\frac{1}{\alpha}}\Big). \tag{2-5}$$

Let us now prove (1-6). In the sequel, we take $0 < \lambda \le \lambda_0$ with λ_0 defined by (2-5). By symmetry, we can assume $|\xi_2| \le |\xi_1|$. We separate different cases:

Case 1: $|\xi_2| \ll |\xi_1|$. Since, by hypothesis, $|\xi_1| \gg 1$, it follows that $\lambda^{-1}|\xi_1| \gg \xi_0$ and thus there exists $\theta \in [\xi_1, \xi_1+\xi_2]$ such that

$$\lambda^{\alpha+1}\big|p(\lambda^{-1}(\xi_1+\xi_2)) - p(\lambda^{-1}\xi_1)\big| = \lambda^{\alpha}|\xi_2||p'(\lambda^{-1}\theta)| \sim \lambda^{\alpha}|\xi_2||\lambda^{-1}\theta|^{\alpha} \sim |\xi_2||\xi_1|^{\alpha}.$$

Now, if $\lambda^{-1}|\xi_2| \le \xi_0$ then

$$\lambda^{\alpha+1}|p(\lambda^{-1}\xi_2)| \le \lambda^{\alpha}|\xi_2| \max_{\xi\in[0,\xi_0]} |p'(\xi)| \ll |\xi_2||\xi_1|^{\alpha}.$$

On the other hand, if $\lambda^{-1}|\xi_2| \ge \xi_0$ then

$$\begin{aligned}\lambda^{\alpha+1}|p(\lambda^{-1}\xi_2)| &= \lambda^{\alpha+1}|p(\xi_0) + p(\lambda^{-1}\xi_2) - p(\xi_0)| \\ &\le \lambda^{\alpha+1}\big(|\xi_0| \max_{\xi\in[0,\xi_0]} |p'(\xi)| + \lambda^{-1}|\xi_2||\lambda^{-1}\xi_2|^{\alpha}\big) \\ &\le |\xi_2|^{\alpha+1} + \lambda^{\alpha}|\xi_2| \max_{\xi\in[0,\xi_0]} |p'(\xi)| \ll |\xi_2||\xi_1|^{\alpha}.\end{aligned}$$

Gathering these two estimates leads to

$$\lambda^{\alpha+1}|\Omega(\lambda^{-1}\xi_1, \lambda^{-1}\xi_2)| \sim |\xi_2||\xi_1|^{\alpha}.$$

Case 2: $|\xi_2| \gtrsim |\xi_1|$. In this case we have $\lambda^{-1}|\xi_2| \gg \xi_0$. Since p is an odd function, by symmetry we can assume that $\xi_1 > 0$.

Case 2(a): $\xi_1\xi_2 \ge 0$. Then we have $0 < \xi_0 \ll \lambda^{-1}\xi_2 \le \lambda^{-1}\xi_1 < \lambda^{-1}\xi_1 + \xi_2$. We notice that

$$\begin{aligned}&\lambda^{\alpha+1}|\Omega(\lambda^{-1}\xi_1, \lambda^{-1}\xi_2)| \\ &\quad = \lambda^{\alpha+1}\int_{\xi_0}^{\lambda^{-1}\xi_2} (p'(\lambda^{-1}\xi_1+\theta) - p'(\theta))\,d\theta + \lambda^{\alpha+1}(p(\lambda^{-1}\xi_1+\xi_0) - p(\lambda^{-1}\xi_1)) - \lambda^{\alpha+1}p(\xi_0)\end{aligned}$$

with

$$|p(\lambda^{-1}\xi_1+\xi_0) - p(\lambda^{-1}\xi_1)| \lesssim \xi_0\lambda^{-\alpha}\xi_1^{\alpha} \ll \lambda^{-\alpha-1}\xi_2\xi_1^{\alpha}$$

and

$$p'(\lambda^{-1}\xi_1+\theta) - p'(\theta) = \int_0^{\lambda^{-1}\xi_1} p''(\theta+\mu)\,d\mu.$$

But, for $\theta \ge \xi_0$, p'' does not change sign since $|p''(\theta)| \sim |\theta|^{\alpha-1}$ and p'' is continuous outside 0. Therefore, for $\theta \in [\xi_0, \lambda^{-1}\xi_2]$, we get

$$\int_0^{\lambda^{-1}\xi_1} p''(\theta+\mu)\,d\mu \sim \int_0^{\lambda^{-1}\xi_1} (\theta+\mu)^{\alpha-1}\,d\mu \sim ((\lambda^{-1}\xi_1+\theta)^{\alpha} - \theta^{\alpha}) \sim \lambda^{-\alpha}\xi_1^{\alpha}.$$

Gathering these estimates we obtain

$$\lambda^{\alpha+1}|\Omega(\lambda^{-1}\xi_1, \lambda^{-1}\xi_2)| \sim \xi_2\xi_1^{\alpha}.$$

Case 2(b): $\xi_1\xi_2 < 0$. For $\xi_1+\xi_2 \ll -\xi_2$, recalling that p is an odd function, we can argue exactly as in Case 1, but with $\xi_1+\xi_2$, $-\xi_2$ and ξ_1 playing the role of ξ_2, ξ_1 and $\xi_1+\xi_2$, respectively. Finally, for $\xi_1+\xi_2 \gtrsim -\xi_2$, we argue exactly as in Case 2(a) with the same exchange of roles as above. □

Lemma 2.2. *Assume that $p_{\alpha+1}$ satisfies (1-6) with $\lambda = 1$. Let $L_1, L_2, L_3 \geq 1$, $0 < N_1 \leq N_2 \leq N_3$ be dyadic numbers and $u, v, w \in \mathcal{S}'(\mathbb{R}^2)$. Then*

$$\int_{\mathbb{R}^2} \Pi(Q_{L_1}P_{N_1}u, Q_{L_2}P_{N_2}v)Q_{L_3}P_{N_3}w = 0$$

whenever the following relation is not satisfied:

$$L_{\max} \sim N_1N_2^\alpha \qquad \text{or} \qquad (L_{\max} \gg N_1N_2^\alpha \quad \text{and} \quad L_{\max} \sim L_{\text{med}}), \tag{2-6}$$

where $L_{\max} = \max(L_1, L_2, L_3)$, $L_{\text{med}} = \max(\{L_1, L_2, L_3\} - \{L_{\max}\})$ and where the two first implicit constants in (2-6) are related to the implicit constant in (1-6).

Proof. This is a direct consequence of the hypothesis (1-6) on the resonance function $\Omega(\xi_1, \xi_2)$, since

$$\Omega(\xi_1, \xi_2) = \sigma(\tau_1+\tau_2, \xi_1+\xi_2) - \sigma(\tau_1, \xi_1) - \sigma(\tau_2, \xi_2)$$

with $\sigma(\tau, \xi) := \tau - p_{\alpha+1}(\xi)$. □

Lemma 2.3. *Let $L \geq 1$, $1 \leq p \leq \infty$ and $s \in \mathbb{R}$. The operator $Q_{\leq L}$ is bounded in $L^p_t H^s$ uniformly in $L \geq 1$.*

Proof. Let $R_{\leq L}$ be the Fourier multiplier by $\eta(\tau/L)$, where η is as defined in Section 2A. The trick is to notice that $Q_{\leq L}u = U_\alpha(t)(R_{\leq L}U_\alpha(-t)u)$. Therefore, using the unitarity of $U_\alpha(\cdot)$ in $H^s(\mathbb{R})$, we infer that

$$\begin{aligned}\|Q_{\leq L}u\|_{L^p_t H^s} = \|U_\alpha(t)(R_{\leq L}U_\alpha(-t)u)\|_{L^p_t H^s} = \|R_{\leq L}U_\alpha(-t)u\|_{L^p_t H^s} \lesssim \|U_\alpha(-t)u\|_{L^p_t H^s} \\ = \|u\|_{L^p_t H^s}.\end{aligned}$$ □

For any $T > 0$, we consider 1_T, the characteristic function of $[0, T]$, and use the decomposition

$$1_T = 1^{\text{low}}_{T,R} + 1^{\text{high}}_{T,R}, \qquad \widehat{1^{\text{low}}_{T,R}}(\tau) = \eta\Big(\frac{\tau}{R}\Big)\widehat{1_T}(\tau) \tag{2-7}$$

for some $R > 0$.

The properties of this decomposition we will need are listed in the following lemmas.

Lemma 2.4. *For any $R > 0$ and $T > 0$,*

$$\|1^{\text{high}}_{T,R}\|_{L^1} \lesssim T \wedge R^{-1} \tag{2-8}$$

and

$$\|1^{\text{low}}_{T,R}\|_{L^\infty} \lesssim 1. \tag{2-9}$$

Proof. A direct computation provides

$$\begin{aligned}
\|1_{T,R}^{\text{high}}\|_{L^1} &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} \Big(1_T(t) - 1_T\Big(t - \frac{s}{R}\Big)\Big) \mathscr{F}^{-1}\eta(s)\, ds \right| dt \\
&\le \int_{\mathbb{R}} \int_{[0,T]\setminus[s/R, T+s/R] \cup [s/R, T+s/R]\setminus[0,T]} |\mathscr{F}^{-1}\eta(s)|\, dt\, ds \\
&\lesssim \int_{\mathbb{R}} \Big(T \wedge \frac{|s|}{R}\Big) |\mathscr{F}^{-1}\eta(s)|\, ds \\
&\lesssim T \wedge R^{-1}.
\end{aligned}$$

Finally, the proof of (2-9) follows directly from the definition of $1_{T,R}^{\text{low}}$ and Young's inequality. □

Lemma 2.5. *Let $u \in L^2(\mathbb{R}^2)$. Then, for any $T > 0$, $R > 0$ and $L \gg R$,*

$$\|Q_L(1_{T,R}^{\text{low}} u)\|_{L^2} \lesssim \|Q_{\sim L} u\|_{L^2}.$$

Proof. By Plancherel we get

$$\begin{aligned}
I_L &= \|Q_L(1_{T,R}^{\text{low}} u)\|_{L^2} \\
&= \|\varphi_L(\tau - \omega(\xi)) \widehat{1_{T,R}^{\text{low}}} *_\tau \hat{u}(\tau, \xi)\|_{L^2} \\
&= \left\| \sum_{L_1 \ge 1} \varphi_L(\tau - \omega(\xi)) \int_{\mathbb{R}} \varphi_{L_1}(\tau' - \omega(\xi)) \hat{u}(\tau', \xi) \eta\Big(\frac{\tau - \tau'}{R}\Big) \frac{e^{-iT(\tau - \tau')} - 1}{\tau - \tau'}\, d\tau' \right\|_{L^2}.
\end{aligned}$$

In the region where $L_1 \ll L$ or $L_1 \gg L$, we have $|\tau - \tau'| \sim L \vee L_1 \gg R$, thus I_L vanishes. On the other hand, for $L \sim L_1$, we get

$$I_L \lesssim \sum_{L \sim L_1} \|Q_L(1_{T,R}^{\text{low}} Q_{L_1} u)\|_{L^2} \lesssim \|Q_{\sim L} u\|_{L^2}.$$ □

3. Unconditional well-posedness in the regular case $s > \frac{1}{2}$

In this section we develop our method in the regular case $s > \frac{1}{2}$. This will emphasize the simplicity of this approach to prove unconditional well-posedness for (1-3) posed on $\mathbb{R}$ or $\mathbb{T}$.

Let $\lambda > 0$ and $L_{\alpha+1}^\lambda$ be the Fourier multiplier by $i\lambda^{\alpha+1} p_{\alpha+1}(\lambda^{-1}\cdot)$. We notice that if u is a solution to (1-3) on $]0, T[$ then $u_\lambda(t, x) = \lambda^\alpha u(\lambda^{\alpha+1} t, \lambda x)$ is a solution to (1-3) on $]0, \lambda^{-(\alpha+1)} T[$ with $L_{\alpha+1}$ replaced by $L_{\alpha+1}^\lambda$. Therefore, up to this change of unknown and equation, we can always assume that the operator $L_{\alpha+1}$ verifies (1-6) with $0 < \lambda \le 1$.

3A. *A priori estimates.* For $s \in \mathbb{R}$ we define the function space M^s as $M^s := L_t^\infty H^s \cap X^{s-1,1}$, endowed with the natural norm

$$\|u\|_{M^s} = \|u\|_{L_t^\infty H^s} + \|u\|_{X^{s-1,1}}.$$

For $u_0 \in H^s(\mathbb{R})$, $s > \frac{1}{2}$, we will construct a solution to (1-3) in M_T^s, whereas the difference of two solutions emanating from initial data belonging to $H^s(\mathbb{R})$ will take place in M_T^{s-1}.

Lemma 3.1. *Let $0 < T < 2$, $s > \frac{1}{2}$ and let $u \in L^\infty_T H^s$ be a solution to (1-3) associated with an initial datum $u_0 \in H^s(\mathbb{R})$. Then $u \in M^s_T$ and*

$$\|u\|_{M^s_T} \lesssim \|u\|_{L^\infty_T H^s} + \|u\|_{L^\infty_T H^s} \|u\|_{L^\infty_T H^{\frac{1}{2}+}}. \tag{3-1}$$

Moreover, for any pair $(u, v) \in (L^\infty_T H^s)^2$ of solutions to (1-3) associated with a pair of initial data $(u_0, v_0) \in (H^s(\mathbb{R}))^2$ and any $s-1 \le r \le s$,

$$\|u-v\|_{M^r_T} \lesssim \|u-v\|_{L^\infty_T H^r} + \|u+v\|_{L^\infty_T H^s} \|u-v\|_{L^\infty_T H^r}. \tag{3-2}$$

Proof. We have to extend the function u from $(0, T)$ to $\mathbb{R}$. For this we follow [Masmoudi and Nakanishi 2005] and introduce the extension operator ρ_T defined by

$$\rho_T u(t) := \eta(t) u\Big(T\mu\Big(\frac{t}{T}\Big)\Big), \tag{3-3}$$

where η is the smooth cut-off function defined in Section 2A and $\mu(t) = \max(1 - |t-1|, 0)$. This ρ_T is a bounded operator from $X^{\theta,b}_T$ into $X^{\theta,b}$ and from $L^p(0, T; X)$ into $L^p(\mathbb{R}; X)$ for any $b \in]-\infty, 1]$, $s \in \mathbb{R}$, $p \in [1, \infty]$ and any Banach space X. Moreover, these bounds are uniform for $0 < T < 1$.

By using this extension operator, it is clear that we only have to estimate the $X^{s-1,1}_T$-norm of u to prove (3-1). As noticed in Remark 1.3, u satisfies the Duhamel formula of (1-3) and $u \in C([0, T]; H^\theta)$ for any $\theta < s$. Hence, standard linear estimates in Bourgain's spaces lead to

$$\begin{aligned}\|u\|_{X^{s-1,1}_T} \lesssim \|u_0\|_{H^{s-1}} + \|\partial_x(u^2)\|_{X^{s-1,0}_T} &\lesssim \|u_0\|_{H^{s-1}} + \|u^2\|_{L^2_T H^s} \\ &\lesssim \|u\|_{L^\infty_T H^{s-1}} + \|u\|_{L^\infty_T H^{\frac{1}{2}+}} \|u\|_{L^\infty_T H^s}\end{aligned}$$

by standard product estimates in Sobolev spaces (see [Adams 1975]).

In the same way, for $s-1 \le r \le s$ we have

$$\|u-v\|_{X^{r-1,1}_T} \lesssim \|u_0-v_0\|_{H^{r-1}} + \|(u+v)(u-v)\|_{L^2_T H^r} \lesssim \|u-v\|_{L^\infty_T H^{r-1}} + \|u+v\|_{L^\infty_T H^s} \|u-v\|_{L^\infty_T H^r},$$

since $s > \frac{1}{2}+$ and $r + s > 0$. This proves (3-2). ☐

Lemma 3.2. *Assume $u_i \in M^0$, $i = 1, 2, 3$, are functions with spatial Fourier support in $\{|\xi| \sim N_i\}$ with $N_i > 0$ dyadic satisfying $N_1 \le N_2 \le N_3$. For any $t > 0$, we set*

$$I_t(u_1, u_2, u_3) = \int_0^t \int_{\mathbb{R}} \Pi(u_1, u_2) u_3.$$

If $N_1 \lesssim 1$,

$$|I_t(u_1, u_2, u_3)| \lesssim N_1^{\frac{1}{2}} \|u_1\|_{L^\infty_t L^2_x} \|u_2\|_{L^2_{tx}} \|u_3\|_{L^2_{tx}}. \tag{3-4}$$

In the case $N_1 \gg 1$,

$$\begin{aligned}|I_t(u_1, u_2, u_3)| \lesssim{}& N_1^{-\frac{1}{2}} N_3^{1-\alpha} \|u_1\|_{L^\infty_t L^2_x} (\|u_2\|_{L^2_{tx}} \|u_3\|_{X^{-1,1}} + \|u_2\|_{X^{-1,1}} \|u_3\|_{L^2_{tx}}) \\ &+ N_1^{\frac{1}{2}} N_3^{-\alpha} \|u_1\|_{X^{-1,1}} \|u_2\|_{L^2_{tx}} \|u_3\|_{L^\infty_t L^2_x} + N_1^{-1} N_3^{-\frac{1}{8}} \|u_1\|_{L^\infty_t L^2_x} \|u_2\|_{L^\infty_t L^2_x} \|u_3\|_{L^\infty_t L^2_x}.\end{aligned}$$

Proof. Estimate (3-4) easily follows from (2-2) together with Hölder's inequality, thus it suffices to estimate $|I_t|$ for $N_1 \gg 1$. Note that I_t vanishes unless $N_2 \sim N_3$. Setting $R = N_1^{3/2} N_3^{1/8}$, we split I_t as

$$I_t(u_1, u_2, u_3) = I_\infty(1_{t,R}^{\mathrm{high}} u_1, u_2, u_3) + I_\infty(1_{t,R}^{\mathrm{low}} u_1, u_2, u_3) := I_t^{\mathrm{high}} + I_t^{\mathrm{low}}, \tag{3-5}$$

where $I_\infty(u, v, w) = \int_{\mathbb{R}^2} \Pi(u, v) w$. The contribution of I_t^{high} is estimated, thanks to Lemma 2.4 as well as (2-2) and Hölder's inequality, by

$$I_t^{\mathrm{high}} \lesssim N_1^{\frac{1}{2}} \|1_{t,R}^{\mathrm{high}}\|_{L^1} \|u_1\|_{L_t^\infty L_x^2} \|u_2\|_{L_t^\infty L_x^2} \|u_3\|_{L_t^\infty L_x^2} \lesssim N_1^{-1} N_3^{-\frac{1}{8}} \|u_1\|_{L_t^\infty L_x^2} \|u_2\|_{L_t^\infty L_x^2} \|u_3\|_{L_t^\infty L_x^2}. \tag{3-6}$$

To evaluate the contribution I_t^{low} we use that, according to Lemma 2.2, we have the decomposition

$$\begin{aligned} I_\infty(1_{t,R}^{\mathrm{low}} u_1, u_2, u_3) = I_\infty(Q_{\gtrsim N_1 N_3^\alpha}(1_{t,R}^{\mathrm{low}} u_1), u_2, u_3) \\ + I_\infty(Q_{\ll N_1 N_2^\alpha}(1_{t,R}^{\mathrm{low}} u_1), Q_{\gtrsim N_1 N_3^\alpha} u_2, u_3) \\ + I_\infty(Q_{\ll N_1 N_2^\alpha}(1_{t,R}^{\mathrm{low}} u_1), Q_{\ll N_1 N_3^\alpha} u_2, Q_{\sim N_1 N_3^\alpha} u_3). \end{aligned} \tag{3-7}$$

It is worth noting that $R \ll N_1 N_3^\alpha$ because $N_1 \gg 1$. Therefore, the contribution $I_t^{1,\mathrm{low}}$ of the first term of the above right-hand side to I_t^{low} is easily estimated, thanks to Lemma 2.5, by

$$I_t^{1,\mathrm{low}} \lesssim N_1^{\frac{1}{2}} (N_1 N_3^\alpha)^{-1} \|u_1\|_{X^{0,1}} \|u_2\|_{L_{tx}^2} \|u_3\|_{L_t^\infty L_x^2} \lesssim N_1^{\frac{1}{2}} N_3^{-\alpha} \|u_1\|_{X^{-1,1}} \|u_2\|_{L_{tx}^2} \|u_3\|_{L_t^\infty L_x^2}. \tag{3-8}$$

Thanks to Lemmas 2.3 and 2.5, the contribution $I_t^{2,\mathrm{low}}$ of the second term can be handled via

$$\begin{aligned} I_t^{2,\mathrm{low}} &\lesssim N_1^{\frac{1}{2}} (N_1 N_3^\alpha)^{-1} \|u_1\|_{L_t^\infty L_x^2} \|u_2\|_{X^{0,1}} \|u_3\|_{L_{tx}^2} \\ &\lesssim N_1^{-\frac{1}{2}} N_3^{1-\alpha} \|u_1\|_{L_t^\infty L_x^2} \|u_2\|_{X^{-1,1}} \|u_3\|_{L_{tx}^2}. \end{aligned} \tag{3-9}$$

Finally, the contribution of the third term is estimated in the same way. □

Remark 3.3. From (2-1) we see that the estimates in Lemma 3.2 also hold for any other rearrangements of N_1, N_2 and N_3.

We are now in position to derive our "improved" energy estimate on smooth solutions to (1-3).

Proposition 3.4. *Let $0 < T < 2$ and let $u \in L_T^\infty H^s$ with $s > \frac{1}{2}$ be a solution to (1-3) associated with an initial datum $u_0 \in H^s(\mathbb{R})$. Then*

$$\|u\|_{L_T^\infty H^s}^2 \lesssim \|u_0\|_{H^s}^2 + (1 + \|u\|_{L_T^\infty H^{\frac{1}{2}+}}^2) \|u\|_{L_T^\infty H^{\frac{1}{2}+}} \|u\|_{L_T^\infty H^s}^2. \tag{3-10}$$

Proof. We apply the operator P_N with $N > 0$ dyadic to (1-3). On account of Remark 1.3, it is clear that $P_N u \in C([0, T]; H^\infty)$ with $\partial_t u_N \in L^\infty(0, T; H^\infty)$. Therefore, taking the L_x^2-scalar product of the resulting equation with $P_N u$, multiplying by $\langle N \rangle^{2s}$ and integrating on $]0, t[$ with $0 < t < T$, we obtain

$$\langle N \rangle^{2s} \|P_N u(t)\|_{L^2}^2 = \langle N \rangle^{2s} \|P_N u_0\|_{L^2}^2 + \langle N \rangle^{2s} \int_0^t \int_{\mathbb{R}} \partial_x P_N(u^2) P_N u.$$

Integrating by parts and applying Bernstein inequalities, this leads to

$$\|P_N u\|^2_{L^\infty_T H^s} \lesssim \|P_N u_0\|^2_{H^s} + \sup_{t\in]0,T[} \langle N\rangle^{2s} \left| \int_0^t \int_{\mathbb{R}} P_N(u^2)\,\partial_x P_N u \right|. \tag{3-11}$$

Thus it remains to estimate

$$I := \sum_{N>0} \langle N\rangle^{2s} \sup_{t\in]0,T[} \left| \int_0^t \int_{\mathbb{R}} P_N(u^2)\,\partial_x P_N u \right|. \tag{3-12}$$

According to (3-1), u belongs to M^s_T. We take an extension $\tilde{u}$ of u supported in time in $]-2,2[$ such that $\|\tilde{u}\|_{M^s} \lesssim \|u\|_{M^s_T}$. To simplify the notation we drop the tilde in the sequel.

By localization considerations, we get

$$P_N(u^2) = P_N(u_{\gtrsim N} u_{\gtrsim N}) + 2P_N(u_{\ll N} u). \tag{3-13}$$

Moreover, using a Taylor expansion of ϕ_N, we easily get

$$P_N(u_{\ll N} u) = u_{\ll N} P_N u + N^{-1}\Pi(\partial_x u_{\ll N}, u), \tag{3-14}$$

where $\Pi = \Pi_\chi$ with $\chi(\xi,\xi_1) = -i\int_0^1 \phi'(N^{-1}(\xi-\theta\xi_1))\,d\theta \in L^\infty$. Inserting (3-13)–(3-14) into (3-12) and integrating by parts, we deduce

$$\begin{aligned} I \lesssim & \sum_{N>0}\sum_{0<N_1\ll N} N_1\langle N\rangle^{2s} \sup_{t\in]0,T[} \left| \int_0^t\int_{\mathbb{R}} \Pi_{\chi_1}(u_{N_1}, u_N) u_N \right| \\ & + \sum_{N>0}\sum_{0<N_1\ll N} N_1\langle N\rangle^{2s} \sup_{t\in]0,T[} \left| \int_0^t\int_{\mathbb{R}} \Pi_{\chi_2}(u_{N_1}, u_{\sim N}) u_N \right| \\ & + \sum_{N>0}\sum_{N_1\gtrsim N} N\langle N\rangle^{2s} \sup_{t\in]0,T[} \left| \int_0^t\int_{\mathbb{R}} \Pi_{\chi_3}(u_{N_1}, u_{\sim N_1}) u_N \right|, \end{aligned}$$

where χ_i, $1\le i\le 3$, are bounded uniformly in N and N_1, and defined by

$$\chi_1(\xi,\xi_1) = \frac{\xi_1}{N_1} 1_{\operatorname{supp}\phi_{N_1}}(\xi_1), \tag{3-15}$$

$$\chi_2(\xi,\xi_1) = \chi(\xi,\xi_1)\frac{\xi_1}{N_1}\frac{\xi}{N}\frac{1_{\operatorname{supp}\phi_N}(\xi)1_{\operatorname{supp}\phi_{N_1}}(\xi_1)}{\phi_{\sim N}(\xi-\xi_1)}, \tag{3-16}$$

$$\chi_3(\xi,\xi_1) = \frac{\xi}{N}\phi_N(\xi). \tag{3-17}$$

Recalling now the definition of I_t (see Lemma 3.2), it follows from (2-1) that

$$I \lesssim \sum_{N>0}\sum_{N_1\gtrsim N} N\langle N_1\rangle^{2s} \sup_{t\in]0,T[} |I_t(u_N, u_{\sim N_1}, u_{N_1})|. \tag{3-18}$$

The contribution of the sum over $N\lesssim 1$ is easily estimated, thanks to (3-4) and Cauchy–Schwarz, by

$$\sum_{N\le 2^9}\sum_{N_1\gtrsim N} N\langle N_1\rangle^{2s} \|u_N\|_{L^\infty_t L^2_x} \|u_{N_1}\|^2_{L^2_t L^2_x} \lesssim \|u\|_{L^\infty_t L^2_x}\|u\|^2_{L^\infty_t H^s}. \tag{3-19}$$

Finally, the contribution of the sum over $N \gg 1$ is controlled with the second part of Lemma 3.2 by

$$\sum_{N>2^9} \sum_{N_1 \gtrsim N} N N_1^{2s} \big[N^{-\frac{1}{2}} N_1^{1-\alpha} \|u_N\|_{L_t^\infty L_x^2} \|u_{N_1}\|_{L^2_{tx}} \|u_{N_1}\|_{X^{-1,1}} + N^{\frac{1}{2}} N_1^{-\alpha} \|u_N\|_{X^{-1,1}} \|u_{N_1}\|^2_{L_t^\infty L_x^2} + N^{-1} N_1^{-\frac{1}{8}} \|u_N\|_{L_t^\infty L_x^2} \|u_{N_1}\|^2_{L_t^\infty L_x^2} \big] \lesssim \|u\|_{M_T^{\frac{1}{2}+}} \|u\|_{M_T^s} \|u\|_{L_T^\infty H^s}. \tag{3-20}$$

Gathering all the above estimates leads to

$$\|u\|^2_{L_T^\infty H^s} \lesssim \|u_0\|^2_{H^s} + \|u\|_{M_T^{\frac{1}{2}+}} \|u\|_{M_T^s} \|u\|_{L_T^\infty H^s}, \tag{3-21}$$

which, together with (3-1), completes the proof of the proposition. □

Let us now establish an a priori estimate at the regularity level $s-1$ on the difference of two solutions.

Proposition 3.5. *Let $0 < T < 2$ and let $u, v \in L_T^\infty H^s$ with $s > \frac{1}{2}$ be two solutions to (1-3) associated with initial data $u_0, v_0 \in H^s(\mathbb{R})$, respectively. Then*

$$\|u - v\|^2_{L_T^\infty H^{s-1}} \lesssim \|u_0 - v_0\|^2_{H^{s-1}} + \|u + v\|_{M_T^s} \|u - v\|^2_{M_T^{s-1}}. \tag{3-22}$$

Proof. The difference $w = u - v$ satisfies

$$w_t + D^\alpha w_x = \partial_x(zw), \tag{3-23}$$

where $z = u + v$. Proceeding as in the proof of Proposition 3.4, we infer that, for $N > 0$,

$$\|P_N w\|^2_{L_T^\infty H^{s-1}} \lesssim \|P_N w_0\|^2_{H^{s-1}} + \sup_{t \in]0,T[} \langle N \rangle^{2(s-1)} \left| \int_0^t \int_{\mathbb{R}} P_N(zw) \partial_x P_N w \right| \tag{3-24}$$

Again, according to (3-1), we can take extensions $\tilde{z}$ and $\widetilde{w}$ of z and w supported in time in $]-2, 2[$ such that $\|\tilde{z}\|_{M^s} \lesssim \|z\|_{M_T^s}$ and $\|\widetilde{w}\|_{M^{s-1}} \lesssim \|w\|_{M_T^{s-1}}$. To simplify the notation we drop the tilde in the sequel.

Setting

$$J := \sum_{N>0} \langle N \rangle^{2(s-1)} \sup_{t \in]0,T[} \left| \int_0^t \int_{\mathbb{R}} P_N(zw) \partial_x P_N w \right|, \tag{3-25}$$

it follows from (3-14) and classical dyadic decomposition that, for all $N > 0$,

$$\begin{aligned} P_N(zw) &= P_N(z_{\ll N} w) + P_N(z_{\sim N} w_{\lesssim N}) + \sum_{N_1 \gg N} P_N(z_{N_1} w_{\sim N_1}) \\ &= z_{\ll N} w_N + N^{-1} \Pi_\chi(\partial_x z_{\ll N}, w) + P_N(z_{\sim N} w_{\lesssim N}) + \sum_{N_1 \gg N} P_N(z_{N_1} w_{\sim N_1}). \end{aligned} \tag{3-26}$$

Inserting this into (3-25) and integrating by parts, we infer

$$
\begin{aligned}
J \lesssim \sum_{N>0}\sum_{N_1\ll N} N_1\langle N\rangle^{2(s-1)}\Big(&\sup_{t\in]0,T[}\Big|\int_0^t\int_{\mathbb{R}}\Pi_{\chi_1}(z_{N_1},w_N)w_N\Big| + \sup_{t\in]0,T[}\Big|\int_0^t\int_{\mathbb{R}}\Pi_{\chi_2}(z_{N_1},w_{\sim N})w_N\Big|\Big)\\
&+\sum_{N>0}\sum_{N_1\lesssim N} N\langle N\rangle^{2(s-1)}\sup_{t\in]0,T[}\Big|\int_0^t\int_{\mathbb{R}}\Pi_{\chi_3}(z_{\sim N},w_{N_1})w_N\Big|\\
&\qquad+\sum_{N>0}\sum_{N_1\gg N} N\langle N\rangle^{2(s-1)}\sup_{t\in]0,T[}\Big|\int_0^t\int_{\mathbb{R}}\Pi_{\chi_3}(z_{N_1},w_{\sim N_1})w_N\Big|,
\end{aligned}
$$

where χ_i, $1\le i\le 3$, are as defined in (3-15)–(3-17). Therefore, it suffices to estimate

$$
\begin{aligned}
J &\lesssim \sum_{N>0}\sum_{N_1\gtrsim N} N\langle N_1\rangle^{2(s-1)}\sup_{t\in]0,T[}|I_t(z_N,w_{\sim N_1},w_{N_1})|\\
&\quad+\sum_{N>0}\sum_{N_1\gtrsim N} N_1\langle N_1\rangle^{2(s-1)}\sup_{t\in]0,T[}|I_t(z_{\sim N_1},w_N,w_{N_1})|\\
&\quad+\sum_{N>0}\sum_{N_1\gtrsim N} N\langle N\rangle^{2(s-1)}\sup_{t\in]0,T[}|I_t(z_{N_1},w_{N_1},w_N)|\\
&:= J_1+J_2+J_3. \qquad (3\text{-}27)
\end{aligned}
$$

The contribution of the sum over $N\lesssim 1$ in (3-27) is easily estimated, thanks to (3-4), by

$$
\begin{aligned}
\sum_{N\lesssim 1}\sum_{N_1\gtrsim N} N^{\frac12}\big(&N\|z_N\|_{L^\infty_tL^2_x}\|w_{N_1}\|^2_{L^2_tH^{s-1}} + N_1\langle N_1\rangle^{-1}\|z_{N_1}\|_{L^2_tH^s}\|w_N\|_{L^\infty_tL^2_x}\|w_{N_1}\|_{L^2_tH^{s-1}}\\
&+N\langle N_1\rangle^{1-2s}\|z_{N_1}\|_{L^2_tH^s}\|w_{N_1}\|_{L^2_tH^{s-1}}\|w_N\|_{L^\infty_tL^2_x}\big)\\
&\lesssim \|z\|_{L^\infty_tL^2_x}\|w\|^2_{L^\infty_tH^{s-1}} + \|w\|_{L^\infty_tH_x^{-\frac12}}\|z\|_{L^\infty_tH^s}\|w\|_{L^\infty_tH^{s-1}}. \qquad (3\text{-}28)
\end{aligned}
$$

For the contribution of the sum over $N\gg 1$, it is worth noting that, since $s>\frac12$, the term J_3 is controlled by J_2. The contribution of J_1 is estimated, thanks to Lemma 3.2, by

$$
\begin{aligned}
\sum_{N\gg 1}\sum_{N_1\gtrsim N} NN_1^{2(s-1)}\big[&N^{-\frac12}N_1^{1-\alpha}\|z_N\|_{L^\infty_tL^2_x}\|w_{N_1}\|_{L^2_{tx}}\|w_{N_1}\|_{X^{-1,1}}\\
&+N^{\frac12}N_1^{-\alpha}\|z_N\|_{X^{-1,1}}\|w_{N_1}\|^2_{L^\infty_tL^2_x} + N^{-1}N_1^{-\frac18}\|z_N\|_{L^\infty_tL^2_x}\|w_{N_1}\|^2_{L^\infty_tL^2_x}\big]\\
&\lesssim \|z\|_{M^{\frac12+}}\|w\|_{M^{s-1}}\|w\|_{L^\infty_tH^{s-1}}. \qquad (3\text{-}29)
\end{aligned}
$$

Finally, in the same way we bound J_2 by

$$
\begin{aligned}
\sum_{N\gg 1}\sum_{N_1\gtrsim N} N_1^{2s-1}\big[&N^{-\frac12}N_1^{1-\alpha}\|w_N\|_{L^\infty_tL^2_x}(\|z_{N_1}\|_{L^2_{tx}}\|w_{N_1}\|_{X^{-1,1}}+\|z_{N_1}\|_{X^{-1,1}}\|w_{N_1}\|_{L^2_{tx}})\\
&+N^{\frac12}N_1^{-\alpha}\|w_N\|_{X^{-1,1}}\|z_{N_1}\|_{L^\infty_tL^2_x}\|w_{N_1}\|_{L^\infty_tL^2_x}\\
&+N^{-1}N_1^{-\frac18}\|w_N\|_{L^\infty_tL^2_x}\|z_{N_1}\|_{L^\infty_tL^2_x}\|w_{N_1}\|_{L^\infty_tL^2_x}\big]\\
&\lesssim \|z\|_{M^s}\|w\|_{M^{-\frac12+}}\|w\|_{L^\infty_tH^{s-1}} + \|z\|_{M^s}\|w\|_{M^{s-1}}\|w\|_{L^\infty_tH^{-\frac12+}}. \qquad (3\text{-}30)
\end{aligned}
$$

Gathering the estimates (3-27)–(3-30), we obtain

$$J \lesssim (\|z\|_{M_T^{\frac{1}{2}+}} \|w\|_{M^{s-1}} + \|z\|_{M_T^s} \|w\|_{M^{-\frac{1}{2}+}}) \|w\|_{L_T^\infty H^{s-1}} + \|z\|_{M_T^s} \|w\|_{M^{s-1}} \|w\|_{L_T^\infty H^{-\frac{1}{2}+}}, \tag{3-31}$$

which leads to (3-22) and completes the proof of the proposition. □

3B. ***Unconditional well-posedness.*** Fix $s > \frac{1}{2}$. First, it is worth noticing that we can always assume that we deal with data that have small H^s-norm. Indeed, if $u \in L^\infty(0, T; H^s)$ is a solution to (1-3), then, for $0 < \lambda \le 1$, $u_\lambda := \lambda^\alpha u(\lambda^{\alpha+1}\cdot, \lambda\cdot) \in L^\infty(0, \lambda^{\alpha+1}T; H^s)$ is a solution to (1-3) with $L_{\alpha+1}$ replaced by $L_{\alpha+1}^\lambda$, that is, the Fourier multiplier by $i\lambda^{\alpha+1} p_{\alpha+1}(\lambda^{-1}\cdot)$. Recall that we assumed at the beginning of this section that $L_{\alpha+1}^\lambda$ satisfies (1-6) for any $0 < \lambda \le 1$. For $0 < \varepsilon \ll 1$, let us denote by $\mathcal{B}^s(\varepsilon)$ the ball of $H^s(\mathbb{R})$ centered at the origin with radius ε. Since

$$\|u_\lambda(0)\|_{H^s} \lesssim \lambda^{\alpha-\frac{1}{2}} \|u_0\|_{H^s},$$

we see that we can force $u_{0,\lambda}$ to belong to $\mathcal{B}^s(\varepsilon)$ by choosing $\lambda = [\varepsilon(1 + \|u_0\|_{H^s})]^{-1/(\alpha-1/2)}$. Therefore, the unconditional well-posedness in $H^s(\mathbb{R})$ of (1-3) for small H^s-initial data with a time of existence $T \ge 1$ will ensure the unconditional well-posedness of (1-3) for arbitrary large H^s-initial data with a maximal time of existence

$$T \gtrsim (1 + \|u_0\|_{H^s})^{-\frac{2(\alpha+1)}{2\alpha-1}}.$$

Existence and unconditional uniqueness. It is well known (see for instance [Abdelouhab et al. 1989]) that (1-3) is locally well-posed in H^s for $s > \frac{3}{2}$ with a minimal time of existence $T = T(\|u_0\|_{H^{3/2+}}) > 0$. So, let $u \in C([0, T_0]; H^\infty(\mathbb{R}))$ be a smooth solution to (1-3) emanating from a smooth initial datum $u_0 \in H^\infty(\mathbb{R})$ with $\|u_0\|_{H^s} \ll 1$. According to (3-10),

$$\|u\|_{L_T^\infty H^s}^2 \lesssim \|u(0)\|_{H^s}^2 + (1 + \|u\|_{L_T^\infty H^{\frac{1}{2}+}}^2) \|u\|_{L_T^\infty H^{\frac{1}{2}+}} \|u\|_{L_T^\infty H^s}^2 \tag{3-32}$$

for any $0 < T \le \min(1, T_0)$ and $s > \frac{1}{2}$. Let us denote by $T^* \ge T_0$ the maximal time of existence of u in $H^\infty(\mathbb{R})$. The well-posedness result in [Abdelouhab et al. 1989] ensures that $\lim_{T \nearrow T^*} \|u\|_{L_T^\infty H^3} = +\infty$ whenever T^* is finite. Since

$$\|u(0)\|_{H^{\frac{1}{2}+}} \le \|u(0)\|_{H^s} \ll 1,$$

(3-32) together with the continuity of $T \mapsto \|u\|_{L_T^\infty H^{1/2+}}$ on $]0, T^*[$ ensure that

$$\|u\|_{L_{T'}^\infty H^{\frac{1}{2}+}} \lesssim \|u(0)\|_{H^{\frac{1}{2}+}} \ll 1$$

with $T' = \min(1, T^*)$. But then (3-32) leads, for any $s > \frac{1}{2}$, to

$$\|u\|_{L_{T'}^\infty H^s} \lesssim \|u(0)\|_{H^s}.$$

This proves that $T' < T^*$ and thus $T' = 1$ and $T^* \ge 1$.

Now, let $u_0 \in H^s(\mathbb{R})$ with $s > \frac{1}{2}$. From the above estimates, we infer that we can pass to the limit on a sequence of solutions $\{u_n\}$ emanating from smooth approximations of u_0 to obtain the existence of a solution $u \in L_T^\infty H^s$ of (1-3) with initial data u_0. Note that one can easily pass to the limit on u_n^2

by compactness arguments, since $\{u_n\}$ and $\{\partial_t u_n\}$ are bounded in $L^\infty_T H^s$ and $L^\infty_T H^{s-3}$, respectively. Estimates (3-22) and (3-1)–(3-2) then ensure that this solution is the only one in this class. Now the continuity of u with values in $H^s(\mathbb{R})$ as well as the continuity of the flow map in $H^s(\mathbb{R})$ will follow from the Bona–Smith argument [1975]. For any $\varphi \in H^s(\mathbb{R})$, dyadic integer $N \geq 1$ and $r \geq 0$, straightforward calculations in Fourier space lead to

$$\|P_{\leq N}\varphi\|_{H^{s+r}_x} \lesssim N^r \|\varphi\|_{H^s_x} \quad \text{and} \quad \|\varphi - P_{\leq N}\varphi\|_{H^{s-r}_x} \lesssim N^{-r}\|P_{>N}\varphi\|_{H^s_x}. \tag{3-33}$$

Let $u_0 \in H^s$ with $s > \frac{1}{2}$ be such that $\|u_0\|_{H^s} \ll 1$. We denote by $u^N \in L^\infty(0,1;H^s)$ the solution of (1-3) emanating from $u_0^N = P_{\leq N}u_0$ and, for $1 \leq N_1 \leq N_2$, we set

$$w := u^{N_1} - u^{N_2}.$$

Then, (3-22) and (3-2) lead to

$$\|w\|_{M_1^{s-1}} \lesssim \|w(0)\|_{H^{s-1}} \lesssim N_1^{-1}\|P_{>N_1}u_0\|_{H^s}. \tag{3-34}$$

Moreover, for any $r \geq 0$ and $s > \frac{1}{2}$, we have

$$\|u^{N_i}\|_{M_1^{s+r}} \lesssim \|u_0^{N_i}\|_{H^{s+r}} \lesssim N_i^r \|u_0\|_{H^s}. \tag{3-35}$$

Next, we observe that w solves the equation

$$w_t + L_{\alpha+1}w = \tfrac{1}{2}\partial_x(w^2) + \partial_x(u^{N_1}w). \tag{3-36}$$

Proposition 3.6. *Let $0 < T < 2$ and let $w \in M^s_T$ with $s > \frac{1}{2}$ be a solution to (3-36). Then*

$$\|w\|^2_{L^\infty_T H^s} \lesssim \|w(0)\|^2_{H^s} + \|w\|^3_{M^s_T} + \|u^{N_1}\|_{M^s_T}\|w\|^2_{M^s_T} + \|u^{N_1}\|_{M^{s+1}_T}\|w\|_{M^s_T}\|w\|_{M^{s-1}_T}. \tag{3-37}$$

Proof. Actually, this is a consequence of estimates derived in the proof of Propositions 3.4 and 3.5. We separate the contributions of $\partial_x(w^2)$ and $\partial_x(u^{N_1}w)$. Let $t \in]0,T[$. First, (3-21) leads to

$$\sum_{N>0} N^{2s}\left|\int_0^t\int_{\mathbb{R}} P_N\partial_x(w^2)P_N w\right| \lesssim \|w\|^3_{M^s_T}.$$

Second, applying (3-31) at the level s with z replaced by u^{N_1}, we obtain

$$\sum_{N>0} N^{2s}\left|\int_0^t\int_{\mathbb{R}} P_N\partial_x(u^{N_1}w)P_N w\right| \lesssim \|u^{N_1}\|_{M^s_T}\|w\|^2_{M^s_T} + \|u^{N_1}\|_{M^{s+1}_T}\|w\|_{M^s_T}\|w\|_{M_T^{-\frac{1}{2}+}},$$

which leads to (3-37) since $s > \frac{1}{2}$. $\square$

Combining (3-2) with (3-37) and (3-35), we get

$$\|w\|^2_{M^s_1} \lesssim (1+\|u_0\|^2_{H^s})\big[\|w_0\|^2_{H^s} + \|u_0\|_{H^s}\|w\|^2_{M^s_1} + \|u_0\|_{H^s}\|w\|^2_{M^s_1} + N_1\|u_0\|_{H^s}\|w\|_{M^s_1}\|w\|_{M^{s-1}_1}\big].$$

Then, the smallness assumption on $\|u_0\|_{H^s}$ and (3-34) lead to

$$\|w\|^2_{M^s_1} \lesssim \|w_0\|^2_{H^s} + N_1^2\|w\|^2_{M^{s-1}_1} \lesssim \|P_{>N_1}u_0\|^2_{H^s}(1+\|P_{>N_1}u_0\|^2_{H^s}) \to 0 \quad \text{as } N_1 \to 0. \tag{3-38}$$

This shows that $\{u^N\}$ is a Cauchy sequence in $C([0,1];H^s)$ and thus $\{u^N\}$ converges in $C([0,1];H^s)$ to a solution of (1-3) emanating from u_0. Then, the uniqueness result ensures that $u\in C([0,1];H^s)$.

Continuity of the flow map. Now let $\{u_{0,n}\}\subset H^s(\mathbb{R})$ be such that $u_{0,n}\to u_0$ in $H^s(\mathbb{R})$. We want to prove that the emanating solution u_n tends to u in $C([0,1];H^s)$. By the triangle inequality, for n large enough,

$$\|u-u_n\|_{L^\infty_1 H^s}\le \|u-u^N\|_{L^\infty_1 H^s}+\|u^N-u^N_n\|_{L^\infty_1 H^s}+\|u^N_n-u_n\|_{L^\infty_1 H^s}.$$

Using the estimate (3-38) on the solution to (3-36) we first infer that

$$\|u-u^N\|_{M^s_1}+\|u_n-u^N_n\|_{M^s_1}\lesssim \|P_{>N}u_0\|_{H^s}+\|P_{>N}u_{0,n}\|_{H^s}$$

and thus

$$\lim_{N\to\infty}\sup_{n\in\mathbb{N}}(\|u-u^N\|_{L^\infty_1 H^s}+\|u_n-u^N_n\|_{L^\infty_1 H^s})=0. \tag{3-39}$$

Next, we notice that (3-22) and (3-2) ensure that

$$\|u^N-u^N_n\|_{M^{s-1}_1}\lesssim \|u^N_0-u^N_{0,n}\|_{H^{s-1}},$$

and thus (3-38) and (3-34) lead to

$$\|u^N-u^N_n\|^2_{M^s_1}\lesssim \|u^N_0-u^N_{0,n}\|^2_{H^s}+N^2\|u^N_0-u^N_{0,n}\|^2_{H^{s-1}}\lesssim \|u_0-u_{0,n}\|^2_{H^s}(1+N^2). \tag{3-40}$$

Combining (3-39) and (3-40), we obtain the continuity of the flow map. The proof of Theorem 1.5 is thus completed in the case $\mathbb{K}=\mathbb{R}$ and $s>\frac{1}{2}$.

3C. *The periodic case.* In this subsection we explain the necessary adaptations to treat the periodic case. First, we define our function spaces in the periodic setting. Since the map $u\mapsto u_\lambda$ maps $L^\infty(0,T;H^s(\mathbb{T}))$ into $L^\infty(0,\lambda^{\alpha+1}T;H^s(\lambda\mathbb{T}))$, we will have to consider space of functions on the tori $\lambda\mathbb{T}$ with $\lambda\ge 1$. We use the same notations as in [Colliander et al. 2004] to deal with Fourier transform of space-periodic functions with a large period $2\pi\lambda$. Then, $(d\xi)_\lambda$ will be the renormalized counting measure on $\lambda^{-1}\mathbb{Z}$:

$$\int a(\xi)\,(d\xi)_\lambda=\frac{1}{\lambda}\sum_{\xi\in\lambda^{-1}\mathbb{Z}}a(\xi)\quad .$$

As noticed in [Colliander et al. 2004], $(d\xi)_\lambda$ is the counting measure on the integers when $\lambda=1$ and converges weakly to the Lebesgue measure when $\lambda\to\infty$. In the definitions below, all the Lebesgue norms in ξ will be with respect to the measure $(d\xi)_\lambda$. For a $2\pi\lambda$-periodic function φ, we define its space Fourier transform on $\lambda^{-1}\mathbb{Z}$ by

$$\hat\varphi(\xi)=\int_{\lambda\mathbb{T}}e^{-i\xi x}f(x)\,dx\quad\text{for all }\xi\in\lambda^{-1}\mathbb{Z}.$$

The Lebesgue spaces $L^q(\lambda\mathbb{T})$, $1\le q\le\infty$, for $2\pi\lambda$-periodic functions, will be defined as usual by

$$\|\varphi\|_{L^q}=\left(\int_{\lambda\mathbb{T}}|\varphi(x)|^q\,dx\right)^{\frac{1}{q}}$$

with the obvious modification for $q=\infty$.

The Sobolev spaces $H^s(\lambda\mathbb{T})$ for $2\pi\lambda$-periodic functions are endowed with the norm

$$\|\varphi\|_{H^s} = \|\langle\xi\rangle^s \hat{\varphi}(\xi)\|_{L^2_\xi} = \|J^s_x \varphi\|_{L^2},$$

where $\langle\cdot\rangle = (1+|\cdot|^2)^{1/2}$ and $\widehat{J^s_x\varphi}(\xi) = \langle\xi\rangle^s\hat{\varphi}(\xi)$.

In the same way, for a function $u(t,x)$ on $\mathbb{R}\times\lambda\mathbb{T}$, we define its space-time Fourier transform by

$$\hat{u}(\tau,\xi) = \mathscr{F}_{t,x}(u)(\tau,\xi) = \int_{\mathbb{R}}\int_{\lambda\mathbb{T}} e^{-i(\tau t+\xi x)} u(t,x)\,dx\,dt \quad \text{for all } (\tau,\xi)\in\mathbb{R}\times\lambda^{-1}\mathbb{Z}.$$

For any $(s,b)\in\mathbb{R}^2$, we define the Bourgain space $X^{s,b}$ of $2\pi\lambda$-periodic (in x) functions as the completion of $\mathscr{S}(\lambda\mathbb{T}\times\mathbb{R})$ for the norm

$$\|u\|_{X^{s,b}} = \|\langle\tau - p_{\alpha+1}(\xi)\rangle^b\langle\xi\rangle^s\hat{u}\|_{L^2_{\tau,\xi}}.$$

Finally, we define the functions ϕ_N and ψ_L and the Fourier multipliers P_N and Q_L as in Section 2A. Since we take a homogeneous decomposition in space frequencies, in the periodic setting

$$u = P_0 u + \sum_{N>0} P_N u, \tag{3-41}$$

where $\widehat{P_0 u}(\xi) = \hat{u}(0)$.

Now, with these definitions, we claim that Lemma 3.1 and Propositions 3.4, 3.5 and 3.6 also hold for $2\pi\lambda$-periodic functions with an implicit constant that does not depend on $\lambda \geq 1$. Indeed, all the tools (the Sobolev and Hölder inequalities) we used in the proofs of these results work also in the periodic setting, independently of the period. However, in view of (3-41), we have to care about the contribution of the null-space frequencies, since we take an homogeneous decomposition. First, since the nonlinear term is a pure derivative, it is clear that the contribution of the null frequency of the nonlinear term vanishes in all the estimates. Now, it is also direct to check that

$$\int_{\lambda\mathbb{T}} P_N(uP_0u)\,\partial_x P_N u = 0 \tag{3-42}$$

and, in the same way,

$$\int_{\lambda\mathbb{T}} P_N(wP_0z)\,\partial_x P_N w = 0. \tag{3-43}$$

We thus just have to control the contribution of the terms $P_N(zP_0w)$ in Proposition 3.5 and $P_N(u^{N_1}P_0w)$ in Proposition 3.6. But the contribution of the first term in Proposition 3.5 can be easily estimated by

$$\begin{aligned} N^{2(s-1)}\left|\int_0^t\int_{\lambda\mathbb{T}} P_N(zP_0w)\,\partial_x P_N w\right| &\lesssim \sup_{t'\in]0,T[} |\hat{w}(t',0)| N^{2(s-1)} N \|P_N z\|_{L^2_T L^2}\|P_N w\|_{L^2_T L^2} \\ &\lesssim \delta_N \|z\|_{L^\infty_T H^s}\|w\|^2_{L^\infty_T H^{s-1}}, \end{aligned}$$

where $\|(\delta_{2^j})_{j\in\mathbb{Z}}\|_{l^1(\mathbb{Z})} \lesssim 1$. Finally, the contribution of the second term in Proposition 3.6 can be estimated in exactly the same way by

$$N^{2s}\left|\int_0^t\int_{\lambda\mathbb{T}} P_N(u^{N_1}P_0w)\,\partial_x P_N w\right| \lesssim \delta_N \|u^{N_1}\|_{L^\infty_T H^{s+1}}\|w\|_{L^\infty_T H^s}\|w\|_{L^\infty_T H^{s-1}}.$$

This completes the proof of the regular case $s > \frac{1}{2}$ in the periodic setting.

4. Estimates in the nonregular case

In this section, we provide the needed estimates at level $s \geq 1-\frac{\alpha}{2}$ for $1<\alpha\leq 2$. We introduce the space

$$F^{s,b} = F^{s,\alpha,b} = X^{s-\frac{\alpha+1}{2},b+\frac{1}{2}} + X^{s-\frac{1+\alpha}{8},b+\frac{1}{8}}, \tag{4-1}$$

endowed with the usual norm, and we define

$$Y^s = Y^{s,\alpha} = L_t^\infty H^s \cap F^{s,\alpha,\frac{1}{2}} = L_t^\infty H^s \cap (X^{s-\frac{\alpha+1}{2},1} + X^{s-\frac{1+\alpha}{8},\frac{5}{8}}).$$

For $u_0 \in H^s(\mathbb{R})$ we will construct a solution to (1-3) that belongs to Y_T^s for some $T = T(\|u_0\|_{H^{1-\alpha/2}})>0$. As in the regular case, by a dilation argument, we may assume that $L_{\alpha+1}$ satisfies (1-6) for $0<\lambda\leq 1$.

Remark 4.1. Except in the case $(s,\alpha)=(0,2)$, we could simply take $Y^{s,\alpha} := L_t^\infty H^s \cap X^{s-(\alpha+1)/2,1}$, since $u \in L^\infty(0,T;H^s)$ forces $\partial_x(u^2) \in L^\infty(0,T;H^{s-(\alpha+1)/2})$. To this point of view, $(s,\alpha)=(0,2)$ is a limit case since $u\in L^\infty(0,T;L^2)$ only implies $\partial_x(u^2)\in L^\infty(0,T;H^{-3/2-})$. As in [Zhou 1997], to overcome this difficulty we have to evaluate our solution in Bourgain's spaces with different conormal regularities.

Lemma 4.2. *Let $0<T<2$, $1<\alpha\leq 2$, $s\geq 1-\frac{\alpha}{2}$ and let $u\in L_T^\infty H^s$ be a solution to (1-3) associated with an initial datum $u_0\in H^s(\mathbb{R})$. Then u belongs to $Y_T^{s,\alpha}$. Moreover, if $(s,\alpha)\neq(0,2)$,*

$$\|u\|_{Y_T^{s,\alpha}} \lesssim \|u\|_{L_T^\infty H^s}(1+\|u\|_{L_T^\infty H^{1-\frac{\alpha}{2}}}) \tag{4-2}$$

and, if $(s,\alpha)=(0,2)$,

$$\|u\|_{Y_T^{0,2}} \lesssim \|u\|_{L_T^\infty L_x^2}(1+\|u\|^2_{L_T^\infty L_x^2}). \tag{4-3}$$

Proof. As in Lemma 3.1 we will work with the extension $\tilde{u}=\rho_T u$ of u (see (3-3)). Recall that $\operatorname{supp}\tilde{u}\subset[-2,2]\times\mathbb{R}$ and that

$$\|\tilde{u}\|_{L_t^\infty H^s} \lesssim \|u\|_{L_T^\infty H^s} \quad \text{and} \quad \|\tilde{u}\|_{X^{\theta,b}} \lesssim \|u\|_{X_T^{\theta,b}}$$

for any $(\theta,b)\in\mathbb{R}\times]-\infty,1]$. It thus remains to control the $F_T^{s,\alpha,\frac{1}{2}}$-norm of u. In the case $(s,\alpha)\neq(0,2)$, we actually simply control the $X_T^{s-(\alpha+1)/2,1}$-norm of u. Using the integral formulation (see Remark 1.3), standard linear estimates in Bourgain's spaces, and standard product estimates in Sobolev spaces, we infer that

$$\begin{aligned}\|u\|_{X_T^{s-\frac{1+\alpha}{2},1}} &\lesssim \|u_0\|_{H^{s-\frac{1+\alpha}{2}}} + \|\partial_x(u^2)\|_{X_T^{s-\frac{1+\alpha}{2},0}} \lesssim \|u_0\|_{H^{s-\frac{1+\alpha}{2}}} + \|u^2\|_{L_T^2 H^{s+\frac{1-\alpha}{2}}} \\ &\lesssim \|u\|_{L_T^\infty H^s} + \|u\|_{L_T^\infty H^{1-\frac{\alpha}{2}}}\|u\|_{L_T^\infty H^s},\end{aligned}$$

since, for $1<\alpha\leq 2$ and $s\geq 1-\frac{\alpha}{2}$ with $(s,\alpha)\neq(0,2)$, we have $s+1-\frac{\alpha}{2}>0$ and $s+1-\frac{\alpha}{2}-(s+\frac{1-\alpha}{2})=\frac{1}{2}$.

Let us now tackle the case $(s,\alpha)=(0,2)$. First we notice that, since $L^1(\mathbb{R})\hookrightarrow H^{-1/2-}(\mathbb{R})$, we have

$$\|u\|_{X_T^{-\frac{7}{4},1}} \lesssim \|u_0\|_{H^{-\frac{7}{4}}} + \|u^2\|_{L_t^2 H^{-\frac{3}{4}}} \lesssim \|u\|_{L_T^\infty L_x^2}(1+\|u\|_{L_T^\infty L_x^2}). \tag{4-4}$$

To bound the $F^{0,2,1/2}$-norm of u, we first notice that linear estimates in Bourgain's spaces lead to

$$\|u\|_{F_T^{0,2,\frac12}} \lesssim \|u_0\|_{H^{-\frac32}} + \|u^2\|_{F_T^{0,2,-\frac12}}$$

and then decompose u^2 as

$$u^2 = P_{\lesssim 1}u^2 + \sum_{N\gg 1}\Big(P_N(P_{\ll N}u u_{\sim N}) + \sum_{N_1'\sim N_1\gtrsim N} P_N(u_{N_1}u_{N_1'})\Big). \tag{4-5}$$

The contribution of the first term in the right-hand side is easily controlled by $\|u\|^2_{L_T^\infty L_x^2}$. The contribution of the second term is easily estimated by

$$\begin{aligned}\Big\|\sum_{N\gg 1}\partial_x P_N(P_{\ll N}u u_{\sim N})\Big\|_{F_T^{0,2,-\frac12}} &\lesssim \Big\|\sum_{N\gg 1}P_N\partial_x(P_{\ll N}u u_{\sim N})\Big\|_{X_T^{-\frac32,0}}\\ &\lesssim \Big(\sum_{N\gg 1}\|P_N(P_{\ll N}u u_{\sim N})\|^2_{L_T^2L_x^1}\Big)^{\frac12}\\ &\lesssim \Big(\sum_{N\geq 1}\|u_N\|^2_{L_T^2L_x^2}\|P_{\ll N}u\|^2_{L_T^\infty L_x^2}\Big)^{\frac12}\\ &\lesssim \|u\|_{L_T^\infty L^2}\|u\|_{L_T^\infty L_x^2}.\end{aligned} \tag{4-6}$$

To estimate the third term, we take advantage of the $X^{-3/8,-3/8}$-part of $F^{0,2,-1/2}$. For $N\gg 1$, we have

$$\begin{aligned}&\sum_{N_1'\sim N_1\gtrsim N}\|\partial_x P_N(P_{N_1}uP_{N_1'}u)\|_{F_T^{0,2,-\frac12}}\\ &\qquad\lesssim \sum_{N_1'\sim N_1\gtrsim N} N\Big\|\sum_{\substack{(L,L_1,L_2)\\ \text{satisfying (2-6)}}}\partial_x P_N Q_L(Q_{L_1}\tilde u_{N_1}Q_{L_2}\tilde u_{N_1'})\Big\|_{X^{-\frac38,-\frac38}}.\end{aligned} \tag{4-7}$$

For the contribution of the sum over $L\gtrsim NN_1^2$ in (4-7), we obtain

$$\begin{aligned}\sum_{N_1\sim N_1'\gtrsim N}\|\partial_x P_N Q_{\gtrsim NN_1^2}(\tilde u_{N_1}\tilde u_{N_1'})\|_{X^{-\frac38,-\frac38}} &\lesssim \sum_{N_1\sim N_1'\gtrsim N} N^{\frac58}N^{\frac12}(NN_1^2)^{-\frac38}\|\tilde u_{N_1}\|_{L^2_{tx}}\|\tilde u_{N_1'}\|_{L_t^\infty L_x^2}\\ &\lesssim \|\tilde u\|_{L_t^\infty L_x^2}\sum_{N_1\gtrsim N}\Big(\frac{N}{N_1}\Big)^{\frac34}\|\tilde u_{N_1}\|_{L^2_{tx}}\\ &\lesssim \gamma_N\|\tilde u\|^2_{L_t^\infty L_x^2}\end{aligned} \tag{4-8}$$

with $\|(\gamma_{2^j})\|_{l^2(\mathbb{N})}\leq 1$. The contribution of the region where $L\ll NN_1^2$ and $L_1\gtrsim NN_1^2$ in (4-7) is controlled by

$$\begin{aligned}&\sum_{N_1\sim N_1'\gtrsim N}\|\partial_x P_N Q_{\ll NN_1^2}(Q_{\gtrsim NN_1^2}\tilde u_{N_1}\tilde u_{N_1'})\|_{X^{-\frac38,-\frac38}}\\ &\quad\lesssim \sum_{N_1\sim N_1'\gtrsim N} N^{\frac58}N^{\frac12}(NN_1^2)^{-1}N_1^{\frac74}\|\tilde u_{N_1}\|_{X^{-\frac74,1}}\|\tilde u_{N_1'}\|_{L_t^\infty L_x^2}\lesssim N^{-\frac18}\|\tilde u\|_{L_t^\infty L_x^2}\|\tilde u\|_{X^{-\frac74,1}}.\end{aligned} \tag{4-9}$$

Finally, the contribution of the last region, where $L, L_1 \ll NN_1^2$ and $L_2 \sim NN_1^2$, in (4-7) is controlled in the same way. Gathering (4-4) and (4-7)–(4-9), we obtain the desired result for the case $(s,\alpha) = (0,2)$. $\square$

In the sequel we will need the following straightforward estimates.

Lemma 4.3. *Let $\alpha \geq 0$ and $w \in F^{0,1/2}$. For $1 \leq B \lesssim N^{\alpha+1}$, we have*

$$\|Q_{\gtrsim B} w_N\|_{L^2} \lesssim B^{-1} N^{\frac{1+\alpha}{2}} \|Q_{\gtrsim B} w_N\|_{F^{0,\frac{1}{2}}} \tag{4-10}$$

and, for $B \gtrsim \langle N\rangle^{\alpha+1}$, we have

$$\|Q_{\gtrsim B} w_N\|_{L^2} \lesssim B^{-\frac{5}{8}} \langle N\rangle^{\frac{1+\alpha}{8}} \|Q_{\gtrsim B} w_N\|_{F^{0,\frac{1}{2}}}. \tag{4-11}$$

Proof. Noticing that $F^{0,1/2} = F^{0,\alpha,1/2} = X^{-(1+\alpha)/2,1} + X^{-(1+\alpha)/8,5/8}$, it is easy to check that

$$\begin{aligned}\|Q_{\gtrsim B} w_N\|_{L^2} &\lesssim \max\big(B^{-1}\langle N\rangle^{\frac{1+\alpha}{2}}, B^{-\frac{5}{8}}\langle N\rangle^{\frac{1+\alpha}{8}}\big)\|Q_{\gtrsim B} w_N\|_{F^{0,\frac{1}{2}}}\\ &\lesssim B^{-\frac{5}{8}}\langle N\rangle^{\frac{1+\alpha}{8}} \max\Big(\Big(\frac{\langle N\rangle^{1+\alpha}}{B}\Big)^{\frac{3}{8}}, 1\Big)\|Q_{\gtrsim B} w_N\|_{F^{0,\frac{1}{2}}},\end{aligned}$$

which leads to the desired result. $\square$

Now we rewrite Lemma 3.2 in the context of the $F^{s,b}$ spaces.

Lemma 4.4. *Assume $u_i \in Y^0$, $i = 1,2,3$, are functions with spatial Fourier support in $\{|\xi| \sim N_i\}$ with $N_i > 0$ dyadic satisfying $N_1 \leq N_2 \leq N_3$.*

If $N_3 \gg 1$ and $N_1 \gtrsim N_3^{2(1-\alpha)/3}$, for $(p,q) \in \{(2,\infty),(\infty,2)\}$,

$$\begin{aligned}|I_t(u_1,u_2,u_3)| \lesssim \sum_{L>1} L^{-1} N_1^{-\frac{1}{2}} N_3^{\frac{1-\alpha}{2}} \|u_1\|_{L_t^p L_x^2} \|Q_{\sim L N_1 N_3^\alpha} u_2\|_{F^{0,\frac{1}{2}}} \|u_3\|_{L_t^q L_x^2}\\ + N_1^{-\frac{1}{2}} N_3^{\frac{1-\alpha}{2}} \|u_1\|_{L_t^p L_x^2} \|u_2\|_{L_t^q L_x^2} \|Q_{\sim N_1 N_3^\alpha} u_3\|_{F^{0,\frac{1}{2}}}\\ + N_1^{-\frac{1}{8}} \langle N_1\rangle^{\frac{1+\alpha}{8}} N_3^{\frac{-5\alpha}{8}} \|u_1\|_{F^{0,\frac{1}{2}}} \|u_2\|_{L_{tx}^2} \|u_3\|_{L_t^\infty L_x^2}\\ + N_1^{-\frac{1}{4}} N_3^{\frac{1}{8}-\frac{\alpha}{2}} \|u_1\|_{L_t^\infty L_x^2} \|u_2\|_{L_t^\infty L_x^2} \|u_3\|_{L_t^\infty L_x^2}.\end{aligned}$$

Proof. For $R = N_1^{3/4} N_3^{\alpha/2-1/8}$, we decompose I_t as in (3-5) and obtain from (3-6) that

$$|I_t^{\text{high}}| \lesssim N_1^{-\frac{1}{4}} N_3^{\frac{1}{8}-\frac{\alpha}{2}} \prod_{i=1}^{3} \|u_i\|_{L_t^\infty L_x^2}.$$

To evaluate I_t^{low} we use the decomposition (3-7) and notice that

$$R = N_1^{\frac{3}{4}} N_3^{\frac{\alpha}{2}-\frac{1}{8}} \leq N_1 N_3^{\frac{2\alpha}{3}-\frac{7}{24}} \ll N_1 N_3^\alpha \quad\text{and}\quad N_1 N_3^\alpha \gtrsim N_3^{\frac{2+\alpha}{3}} \gg 1.$$

Therefore, the contribution $I_t^{1,\text{low}}$ of the first term of the right-hand side of (3-7) to I_t^{low} is easily estimated, thanks to Lemmas 2.5 and 4.3, by

$$|I_t^{1,\text{low}}| \lesssim N_1^{\frac{1}{2}} (N_1 N_3^\alpha)^{-\frac{5}{8}} \langle N_1\rangle^{\frac{\alpha+1}{8}} \|u_1\|_{F^{0,\frac{1}{2}}} \|u_2\|_{L_{tx}^2} \|u_3\|_{L_t^\infty L_x^2},$$

which is acceptable. Thanks to Lemmas 2.3, 2.5 and 4.3, the contribution $I_t^{2,\mathrm{low}}$ of the second term can be handled in the following way:

$$
\begin{aligned}
|I_t^{2,\mathrm{low}}| &\lesssim \sum_{L>1} N_1^{\frac12}(LN_1N_3^\alpha)^{-1}N_3^{\frac{\alpha+1}{2}}\|u_1\|_{L_t^pL_x^2}\|Q_{\sim LN_1N_3^\alpha}u_2\|_{F^{0,\frac12}}\|u_3\|_{L_t^qL_x^2}\\
&\lesssim \sum_{L>1} L^{-1}N_1^{-\frac12}N_3^{\frac{1-\alpha}{2}}\|u_1\|_{L_t^pL_x^2}\|Q_{\sim LN_1N_3^\alpha}u_2\|_{F^{0,\frac12}}\|u_3\|_{L_t^qL_x^2}.
\end{aligned}
\tag{4-12}
$$

In the same way, we get that the contribution $I_t^{3,\mathrm{low}}$ of the third term in I_t^{low} is bounded by

$$
\begin{aligned}
|I_t^{3,\mathrm{low}}| &\lesssim N_1^{\frac12}(N_1N_3^\alpha)^{-1}N_3^{\frac{\alpha+1}{2}}\|u_1\|_{L_t^pL_x^2}\|u_2\|_{L_t^qL_x^2}\|Q_{\sim N_1N_3^\alpha}u_3\|_{F^{0,\frac12}}\\
&\lesssim N_1^{-\frac12}N_3^{\frac{1-\alpha}{2}}\|u_1\|_{L_t^pL_x^2}\|u_2\|_{L_t^qL_x^2}\|Q_{\sim N_1N_3^\alpha}u_3\|_{F^{0,\frac12}}.
\end{aligned}
\tag{4-13}
$$

Gathering all these estimates, we obtain the desired bound. □

Proposition 4.5. *Let $0<T<2$, $1<\alpha\le 2$, $s\ge 1-\frac{\alpha}{2}$ and let $u\in L_T^\infty H^s$ be a solution to (1-3) associated with an initial datum $u_0\in H^s(\mathbb{R})$. Then u belongs to $\widetilde{L}_T^\infty H^s$ and*

$$
\|u\|_{\widetilde{L}_T^\infty H^s}^2 \lesssim \|u_0\|_{H^s}^2 + \|u\|_{L_T^\infty H^s}\big(\|u\|_{L_T^\infty H^{1-\frac{\alpha}{2}}}\|u\|_{Y_T^s} + \|u\|_{L_T^\infty H^s}\|u\|_{Y_T^{1-\frac{\alpha}{2}}}\big).
\tag{4-14}
$$

Proof. First, we notice that Lemma 4.2 ensures that $u\in Y_T^s$. Applying the operator P_N with $N>0$ dyadic to (1-3), arguing as in (3-11), we obtain

$$
\|P_Nu\|_{L_T^\infty H^s}^2 \lesssim \|P_Nu_0\|_{H^s}^2 + \sup_{t\in]0,T[}\langle N\rangle^{2s}\left|\int_0^t\int_{\mathbb{R}} P_N(u^2)\,\partial_x P_N u\right|.
\tag{4-15}
$$

We take an extension $\tilde u$ of u supported in time in $]-4,4[$ such that $\|\tilde u\|_{Y^s}\lesssim\|u\|_{Y_T^s}$. To simplify the notation we drop the tilde in the sequel. We infer from (3-18) that it suffices to estimate

$$
I=\sum_{N>0}\sum_{N_1\gtrsim N} N\langle N_1\rangle^{2s}\sup_{t\in]0,T[}|I_t(u_N,u_{\sim N_1},u_{N_1})|.
$$

The low frequencies part, $N\lesssim 1$, is estimated exactly as in (3-19) by

$$
\|u\|_{L_t^\infty L_x^2}\|u\|_{L_t^\infty H^s}^2.
$$

On the other hand, the contribution of the sum over $N\gg 1$ is controlled, thanks to Lemma 4.4, by

$$
\begin{aligned}
\sum_{N\gg1}\sum_{N_1\gtrsim N}\Big[\Big(\frac{N}{N_1}\Big)^{\frac{\alpha-1}{2}}&\|u_N\|_{L_t^2H^{1-\frac{\alpha}{2}}}\|u_{N_1}\|_{L_t^\infty H^s}\|u_{N_1}\|_{F^{s,\frac12}}\\
&+\Big(\frac{N}{N_1}\Big)^{\frac{5\alpha}{8}}\|u_N\|_{F^{1-\frac{\alpha}{2},\frac12}}\|u_{N_1}\|_{L_t^2H^s}\|u_{N_1}\|_{L_t^\infty H^s}\\
&\qquad+N^{\frac{\alpha}{2}-\frac14}N_1^{\frac18-\frac{\alpha}{2}}\|u_N\|_{L_t^\infty H^{1-\frac{\alpha}{2}}}\|u_{N_1}\|_{L_t^\infty H^s}^2\Big]\\
&\lesssim \|u\|_{Y^{1-\frac{\alpha}{2}}}\|u\|_{L_t^\infty H^s}^2+\|u\|_{L_t^\infty H^{1-\frac{\alpha}{2}}}\|u\|_{L_t^\infty H^s}\|u\|_{Y^s},
\end{aligned}
\tag{4-16}
$$

where we use the discrete Young's inequality in N_1 and then Cauchy–Schwarz in N to bound the first two terms.

Gathering the above estimates we eventually obtain

$$I \lesssim \|u\|_{Y_T^{1-\frac{\alpha}{2}}} \|u\|_{L_T^\infty H^s}^2 + \|u\|_{L_T^\infty H^{1-\frac{\alpha}{2}}} \|u\|_{L_T^\infty H^s} \|u\|_{Y_T^s}, \tag{4-17}$$

which completes the proof of the proposition. □

4A. ***Estimates on the difference of two solutions.*** First we introduce the function spaces where we will estimate the difference of two solutions of (1-3). Contrary to the regular case, we will have to work in a function space that puts a weight on the very low frequencies. This kind of weighted space for the difference of two solutions was, for instance, used in [Ionescu et al. 2008] in the context of short-time Bourgain spaces.

For $\theta \in \mathbb{R}$ we define the Banach space

$$\overline{H}^\theta(\mathbb{R}) = \{\varphi \in H^\theta(\mathbb{R}) \mid \|\varphi\|_{\overline{H}^\theta} < \infty\}$$

with

$$\|\varphi\|_{\overline{H}^\theta} := \|\langle |\xi|^{-\frac{1}{2}}\rangle \langle \xi\rangle^\theta \hat{\varphi}\|_{L^2},$$

equipped with the norm $\|\cdot\|_{\overline{H}^\theta}$. Then we define the space $\widetilde{L}_t^\infty \overline{H}^\theta$ by

$$\|w\|_{\widetilde{L}_t^\infty \overline{H}^\theta} := \left(\sum_{N>0} \|w_N\|_{L_t^\infty \overline{H}^\theta}^2\right)^{\frac{1}{2}}. \tag{4-18}$$

Finally, we define the function spaces $\widetilde{Y}^\theta$ and Z^θ, $\theta \in \mathbb{R}$, by

$$\widetilde{Y}^\theta = \widetilde{L}_t^\infty H^\theta \cap F^{\theta,\frac{1}{2}} \quad \text{and} \quad Z^\theta = \widetilde{L}_t^\infty \overline{H}^\theta \cap F^{\theta,\frac{1}{2}},$$

with $F^{\theta,b}$ as defined in (4-1).

If $u, v \in L_T^\infty H^s$ are two solutions of (1-3) with $s \geq 1-\frac{\alpha}{2}$, then, by Lemma 4.2 and Proposition 4.5, we know that u and v belong to $Y_T^s \cap \widetilde{L}_T^\infty H^s$. Moreover, again using the extension operator ρ_T, it is easy to check that

$$Y_T^s \cap \widetilde{L}_T^\infty H^s \hookrightarrow \widetilde{Y}_T^s \tag{4-19}$$

with an embedding constant that does not depend on $0 < T \leq 2$. Hence, u and v belong to $\widetilde{Y}_T^s$. Assuming that $u_0 - v_0 \in \overline{H}^s$, we claim that the difference $u-v$ belongs to Z_T^s. Indeed, according to the above definitions of $\widetilde{Y}^s$ and Z^s, it suffices to check that $P_1(u-v)$ belongs to $\widetilde{L}_T^\infty \overline{H}^s$. But this is straightforward, since, by the Duhamel formula, for any dyadic integer $0 < N < 1$ we have

$$\|P_N(u-v)\|_{L_T^\infty \overline{H}^s} \lesssim \|u_0 - v_0\|_{\overline{H}^s} + N^{\frac{1}{2}}(\|u\|_{L_T^\infty L_x^2}^2 + \|v\|_{L_T^\infty L_x^2}^2).$$

We are thus allowed to estimate the difference $w = u - v$ in the space $Z_T^{s-3/2+\alpha/2}$.

Remark 4.6. For $\alpha > 1$, we have $s - \frac{3}{2} + \frac{\alpha}{2} > s - 1$ and thus, contrary to the preceding section, the derivative of a solution does not belong to the space where we estimate the difference $w = u - v$ of two solutions. This fact is crucial in the preceding section to recover the derivative in terms as J_2 in (3-27) that contains small space frequencies of w. In this section, we will instead combine the weight on the low space frequencies of w with the resonance relation to control such contributions.

Proposition 4.7. *Let $0 < T < 2$, $1 < \alpha \le 2$, $s \ge 1 - \frac{\alpha}{2}$ and $u, v \in L^\infty_T H^s$ be two solutions to (1-3) on $]0, T[$ associated with initial data $u_0, v_0 \in H^s$ such that $u_0 - v_0 \in \overline{H}^s$. Then $u - v \in Z_T^{s-3/2+\alpha/2}$ and we have*

$$\|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}} \lesssim \|u-v\|_{L_T^\infty \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|u+v\|_{\widetilde{Y}_T^s} \|u-v\|_{Z_T^{-\frac{1}{2}}} + \|u+v\|_{\widetilde{Y}^{1-\frac{\alpha}{2}}} \|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}. \tag{4-20}$$

Proof. The fact that $u - v \in Z_T^{s-3/2+\alpha/2}$ follows from the discussion above. Now, recall that $w = u - v$ satisfies (3-23) with $z = u + v$. We extend w from $(0, T)$ to $\mathbb{R}$ by using the extension operator ρ_T defined in (3-3). On account of the uniform bounds on ρ_T (see the paragraph just after (3-3)), it remains to estimate the $F_T^{s-3/2+\alpha/2,\alpha,1/2}$-norm of w. From classical linear estimates in the framework of Bourgain's spaces, the Duhamel formulation associated with (3-23) leads to

$$\|w\|_{F_T^{s-\frac{3}{2}+\frac{\alpha}{2},\frac{1}{2}}} \lesssim \|w_0\|_{H^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|\partial_x(zw)\|_{F_T^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}}. \tag{4-21}$$

Let $\tilde{z}$ and $\widetilde{w}$ be time extensions of z and w satisfying $\|\tilde{z}\|_{\widetilde{Y}^s} \lesssim \|z\|_{\widetilde{Y}_T^s}$ and $\|\widetilde{w}\|_{Z^{s-3/2+\alpha/2}} \lesssim \|w\|_{Z_T^{s-3/2+\alpha/2}}$. To simplify the notation we drop the tilde in the sequel. From (4-21) we see that it suffices to estimate

$$\|\partial_x(zw)\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} \lesssim \Bigl(\sum_{N>0} \|P_N \partial_x(zw)\|^2_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}}\Bigr)^{\frac{1}{2}}.$$

We first estimate the low-high contribution $P_N(P_{\lesssim N} z P_{\sim N} w)$:

$$\begin{aligned}
\|\partial_x P_N(P_{\lesssim N} z P_{\sim N} w)\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} &\lesssim \sum_{N_1 \lesssim N} N \|P_N(P_{N_1} z P_{\sim N} w)\|_{X^{s-2,0}} \\
&\lesssim \sum_{N_1 \lesssim N} N_1^{\frac{1}{2}} N \langle N\rangle^{s-2} \|P_{N_1} z\|_{L_t^\infty L_x^2} \|P_{\sim N} w\|_{L_t^2 L_x^2} \\
&\lesssim \|P_{\sim N} w\|_{L_t^2 H^{s-\frac{3}{2}+\frac{\alpha}{2}}} \sum_{N_1 \lesssim N} \Bigl(\frac{N_1}{\langle N\rangle}\Bigr)^{\frac{\alpha-1}{2}} \|P_{N_1} z\|_{L_t^\infty H^{1-\frac{\alpha}{2}}} \\
&\lesssim \|z\|_{L_t^\infty H^{1-\frac{\alpha}{2}}} \|P_{\sim N} w\|_{L_t^\infty H^{s-\frac{3}{2}+\frac{\alpha}{2}}}.
\end{aligned}$$

Similarly, the high-low interactions are estimated as follows:

$$\begin{aligned}
\|\partial_x P_N(P_{\sim N} z P_{\lesssim N} w)\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} &\lesssim N \|P_N(P_{\sim N} z P_{\lesssim N} w)\|_{X^{s-2,0}} \\
&\lesssim \|P_{\sim N} z\|_{L_t^2 H^s} \sum_{N_1 \lesssim N} \Bigl(\frac{N_1}{\langle N\rangle}\Bigr)^{\frac{1}{2}} \|P_{N_1} w\|_{L_t^\infty H^{-\frac{1}{2}}} \\
&\lesssim \|P_{\sim N} z\|_{L_t^2 H^s} \|w\|_{L_t^\infty H^{-\frac{1}{2}}}.
\end{aligned}$$

Now we deal with the high-high interactions term:

$$\|\partial_x P_N(P_{\gg N} z P_{\gg N} w)\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} \lesssim \sum_{N_1 \gg N} N \Big\| \sum_{\substack{(L,L_1,L_2)\\ \text{satisfying (2-6)}}} \partial_x P_N Q_L(Q_{L_1} z_{N_1} Q_{L_2} w_{N_1}) \Big\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}}.$$

We may assume that $N_1 \gg 1$ since, otherwise, $N \ll N_1 \lesssim 1$ and we have

$$\|P_{\lesssim 1}\partial_x(P_{\lesssim 1} z P_{\lesssim 1} w)\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} \lesssim \|P_{\lesssim 1} z\|_{L_t^\infty L^2} \|P_{\lesssim 1} w\|_{L_t^\infty H^{-\frac{1}{2}}}.$$

For $N_1 \gg 1$, we will take advantage of the fact that $X^{s-13/8+3\alpha/8,-3/8} \hookrightarrow F^{s-3/2+\alpha/2,-1/2}$. The contribution of the sum over $L \gtrsim N N_1^\alpha$ can be thus controlled by

$$\begin{aligned}
&\sum_{N_1 \gg N} \|\partial_x P_N Q_{\gtrsim N N_1^\alpha}(z_{N_1} w_{N_1})\|_{F^{s-\frac{3}{2}+\frac{\alpha}{2},-\frac{1}{2}}} \\
&\qquad \lesssim \sum_{N_1 \gg N} N \|P_N Q_{\gtrsim N N_1^\alpha}(z_{N_1} w_{N_1})\|_{X^{s-\frac{13}{8}+\frac{3\alpha}{8},-\frac{3}{8}}} \\
&\qquad \lesssim \sum_{N_1 \gg N} \sum_{L \gtrsim N N_1^\alpha} N \langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} L^{-\frac{3}{8}} \|P_N Q_L(z_{N_1} w_{N_1})\|_{L^2} \\
&\qquad \lesssim \sum_{N_1 \gg N} N^{\frac{3}{2}} \langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} (N N_1^\alpha)^{-\frac{3}{8}} N_1^{\frac{1}{2}-s} \|z_{N_1}\|_{L_t^2 H^s} \|w_{N_1}\|_{L_t^\infty H^{-\frac{1}{2}}} \\
&\qquad \lesssim \sum_{N_1 \gg N} \left(\frac{N}{N_1}\right)^{\frac{1}{2}-\frac{\alpha}{8}} \left(\frac{\langle N\rangle}{\langle N_1\rangle}\right)^{s-1+\frac{\alpha}{2}} \|z_{N_1}\|_{L_t^2 H^s} \|w_{N_1}\|_{L_t^\infty H^{-\frac{1}{2}}} \\
&\qquad \lesssim \delta_N \|z\|_{L_t^2 H^s} \|w\|_{L_t^\infty H^{-\frac{1}{2}}},
\end{aligned}$$

where $\|(\delta_{2^j})_j\|_{l^2(\mathbb{Z})} \lesssim 1$. The contribution of the region where $L \ll N N_1^\alpha$ and $L_1 \gtrsim N N_1^\alpha$ is estimated, thanks to (4-10), by

$$\begin{aligned}
&\sum_{N_1 \gg N} \|\partial_x P_N Q_{\ll N N_1^\alpha}(Q_{\gtrsim N N_1^\alpha} z_{N_1} w_{N_1})\|_{X^{s-\frac{13}{8}+\frac{3\alpha}{8},-\frac{3}{8}}} \\
&\qquad \lesssim \sum_{N_1 \gg N} N \langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} \|P_N(Q_{\gtrsim N N_1^\alpha} z_{N_1} w_{N_1})\|_{L^2} \\
&\qquad \lesssim \sum_{N_1 \gg N} N^{\frac{3}{2}} \langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} (N N_1^\alpha)^{-1} N_1^{1-s+\frac{\alpha}{2}} \|Q_{\gtrsim N N_1^\alpha} z_{N_1}\|_{F^{s,\frac{1}{2}}} \|w_{N_1}\|_{L_t^\infty H^{-\frac{1}{2}}} \\
&\qquad \lesssim \sum_{N_1 \gg N} \left(\frac{N}{\langle N\rangle}\right)^{\frac{1}{2}} \langle N\rangle^{-\frac{1+\alpha}{8}} \left(\frac{\langle N\rangle}{\langle N_1\rangle}\right)^{s-1+\frac{\alpha}{2}} \|Q_{\gtrsim N N_1^\alpha} z_{N_1}\|_{F^{s,\frac{1}{2}}} \|w_{N_1}\|_{L_t^\infty H^{-\frac{1}{2}}} \\
&\qquad \lesssim \delta_N \|z\|_{Y^s} \|w\|_{\widetilde{L}_t^\infty H^{-\frac{1}{2}}},
\end{aligned}$$

where $\|(\delta_{2^j})_j\|_{l^2(\mathbb{Z})} \lesssim 1$. Finally the contribution of the last region can be bounded, thanks to (4-10), by

$$\begin{aligned}
&\sum_{N_1 \gg N} \|\partial_x P_N Q_{\ll NN_1^\alpha}(Q_{\ll NN_1^\alpha} z N_1 Q_{\sim NN_1^\alpha} w_{N_1})\|_{X^{s-\frac{13}{8}+\frac{3\alpha}{8},-\frac{3}{8}}} \\
&\quad \lesssim \sum_{N_1 \gg N} N\langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} \|P_N Q_{\ll NN_1^\alpha}(Q_{\ll NN_1^\alpha} z_{N_1} Q_{\sim NN_1^\alpha} w_{N_1})\|_{L^2} \\
&\quad \lesssim \sum_{N_1 \gg N} N^{\frac{3}{2}}\langle N\rangle^{s-\frac{13}{8}+\frac{3\alpha}{8}} N_1^{-s}(NN_1^\alpha)^{-1} N_1^{1+\frac{\alpha}{2}} \|Q_{\ll NN_1^\alpha} z_{N_1}\|_{L_t^\infty H^s} \|Q_{\sim NN_1^\alpha} w_{N_1}\|_{F^{-\frac{1}{2},\frac{1}{2}}} \\
&\quad \lesssim \sum_{N_1 \gg N} \left(\frac{N}{\langle N\rangle}\right)^{\frac{1}{2}} \langle N\rangle^{-\frac{1+\alpha}{8}} \left(\frac{\langle N\rangle}{\langle N_1\rangle}\right)^{s-1+\frac{\alpha}{2}} \|z_{N_1}\|_{L_t^\infty H^s} \|w_{N_1}\|_{F^{-\frac{1}{2},\frac{1}{2}}} \\
&\quad \lesssim \delta_N \|z\|_{\widetilde{L}_t^\infty H^s} \|w\|_{Z^{-\frac{1}{2}}},
\end{aligned}$$

which is acceptable. This concludes the proof of Proposition 4.7. ☐

Proposition 4.8. *Let $1 \le \alpha \le 2$, $0 < T < 2$ and let $u, v \in L_T^\infty H^s$ with $s \ge 1 - \frac{\alpha}{2}$ be two solutions to (1-3) associated with initial data $u_0, v_0 \in H^s$ such that $u_0 - v_0 \in \overline{H}^s$. Then*[1]

$$\|u-v\|^2_{\widetilde{L}_T^\infty \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} \lesssim \|u_0-v_0\|^2_{\overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|u+v\|_{Y_T^s} \|u-v\|_{\widetilde{L}_T^\infty \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} \|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}. \tag{4-22}$$

Proof. Recall that the difference $w = u - v$ satisfies (3-23) with $z = u + v$. Applying the operator P_N with $N > 0$ dyadic to (3-23), taking the L^2 scalar product with $P_N w$ and integrating on $]0, t[$, we obtain

$$\|w_N\|^2_{L_T^\infty \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} \lesssim \|P_N w_0\|_{\overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \langle N^{-1}\rangle\langle N\rangle^{2(s-\frac{3}{2}+\frac{\alpha}{2})} \sup_{t\in[0,T]} \left|\int_0^t\int_{\mathbb{R}} P_N(zw)\partial_x w_N\right|.$$

Therefore, we have to estimate

$$J := \sum_{N>0} \langle N^{-1}\rangle\langle N\rangle^{2(s-\frac{3-\alpha}{2})} \sup_{t\in[0,T]} \left|\int_0^t\int_{\mathbb{R}} P_N(zw)\partial_x w_N\right|.$$

We take extensions $\tilde{z}$ and $\widetilde{w}$ of z and w supported in time in $]-4, 4[$ such that $\|\tilde{z}\|_{Y^s} \lesssim \|u\|_{Y_T^s}$ and $\|\widetilde{w}\|_{Z^s} \lesssim \|w\|_{Z_T^s}$. To simplify the notation we drop the tilde in the sequel.

Proceeding as in (3-27), we get

$$\begin{aligned}
J &\lesssim \sum_{N>0}\sum_{N_1 \gtrsim N} N\langle N_1^{-1}\rangle\langle N_1\rangle^{2(s-\frac{3-\alpha}{2})} \sup_{t\in]0,T[} |I_t(z_N, w_{\sim N_1}, w_{N_1})| \\
&\quad + \sum_{N>0}\sum_{N_1 \gtrsim N} N_1\langle N_1^{-1}\rangle\langle N_1\rangle^{2(s-\frac{3-\alpha}{2})} \sup_{t\in]0,T[} |I_t(z_{\sim N_1}, w_N, w_{N_1})| \\
&\quad\quad + \sum_{N>0}\sum_{N_1 \gtrsim N} N\langle N^{-1}\rangle\langle N\rangle^{2(s-\frac{3-\alpha}{2})} \sup_{t\in]0,T[} |I_t(z_{N_1}, w_{N_1}, w_N)| \\
&:= J_1 + J_2 + J_3.
\end{aligned} \tag{4-23}$$

[1] We include the case $\alpha = 1$ here since it does not lead to additional difficulties and will be useful in the Appendix to prove LWP for $(\alpha, s) = (1, \frac{1}{2})$.

Estimates for J_1. The contribution of the sum over $N \lesssim 1$ in J_1 is estimated, thanks to (3-4), by

$$\sum_{N\lesssim 1}\sum_{N_1\gtrsim N} N^{\frac{3}{2}}\|z_N\|_{L_t^\infty L_x^2}\|w_{N_1}\|^2_{L_t^\infty \overline{H}^{s-\frac{3-\alpha}{2}}} \lesssim \|z\|_{L_t^\infty L_x^2}\|w\|^2_{\widetilde{L}_T^\infty \overline{H}^{s-\frac{3-\alpha}{2}}}.$$

The contribution $N \gg 1$ in J_1 can be controlled with Lemma 4.4 by

$$\begin{aligned}\sum_{N\gg 1}\sum_{N_1\gtrsim N}\Bigg[\sum_{L>1} L^{-1}\Big(\frac{N}{N_1}\Big)^{\frac{\alpha-1}{2}}\|z_N\|_{L_t^2 H^{1-\frac{\alpha}{2}}}\|Q_{\sim LNN_1^\alpha} w_{N_1}\|_{F^{s-\frac{3-\alpha}{2},\frac12}}\|w_{N_1}\|_{L_t^\infty H^{s-\frac{3-\alpha}{2}}}\\ + \Big(\frac{N}{N_1}\Big)^{\frac{5\alpha}{8}}\|z_N\|_{F^{1-\frac{\alpha}{2},\frac12}}\|w_{N_1}\|_{L_t^2 H^{s-\frac{3-\alpha}{2}}}\|w_{N_1}\|_{L_t^\infty H^{s-\frac{3-\alpha}{2}}}\\ + N^{\frac{\alpha}{2}-\frac14}N_1^{\frac18-\frac{\alpha}{2}}\|z_N\|_{L_t^\infty H^{1-\frac{\alpha}{2}}}\|w_{N_1}\|^2_{L_t^\infty H^{s-\frac{3-\alpha}{2}}}\Bigg]\\ \lesssim \|z\|_{Y^{1-\frac{\alpha}{2}}}\|w\|_{\widetilde{L}_t^\infty \overline{H}^{s-\frac{3-\alpha}{2}}}\|w\|_{Z^{s-\frac{3-\alpha}{2}}},\end{aligned}$$

where for the first term we used Cauchy–Schwarz in (N, N_1) and then summed in L. Note that for $\alpha > 1$ we could replace the $\widetilde{L}_t^\infty H^{s-3/2+\alpha/2}$-norm by a standard $L_t^\infty H^{s-3/2+\alpha/2}$-norm by invoking the discrete Young inequality.

Estimates for J_2. We separate different contributions. First, the contribution of the sum over $N_1 \lesssim 1$ is directly estimated by $\|z\|_{L_T^\infty L^2}\|w\|^2_{L_T^\infty H^{-1/2}}$. The contribution of the sum over $N \le N_1^{2(1-\alpha)/3}$ and $N_1 \gg 1$ is then easily estimated by

$$\begin{aligned}\sum_{N_1\gg 1}\sum_{N\le N_1^{\frac23(1-\alpha)}} NN_1^{\frac{\alpha-1}{2}}\|z_{N_1}\|_{L_T^2 H^s}\|w_N\|_{L_T^\infty \overline{H}^{-\frac12}}\|w_{N_1}\|_{L_T^2 H^{s-\frac32+\frac{\alpha}{2}}}\\ \lesssim \sum_{N_1\gg 1} N_1^{\frac{1-\alpha}{6}}\|z_{N_1}\|_{L_T^2 H^s}\|w\|_{L_T^\infty \overline{H}^{-\frac12}}\|w_{N_1}\|_{L_T^2 H^{s-\frac32+\frac{\alpha}{2}}}\\ \lesssim \|z\|_{L_T^\infty H^s}\|w\|_{L_T^\infty \overline{H}^{-\frac12}}\|w\|_{L_T^\infty H^{s-\frac32+\frac{\alpha}{2}}}.\end{aligned} \tag{4-24}$$

Finally the contribution of the sum over $N_1 \gg 1$ and $N \gg N_1^{2(1-\alpha)/3}$ is bounded, thanks to Lemma 4.4, by

$$\begin{aligned}\sum_{N_1\gg 1}\sum_{N\gg N_1^{\frac23(1-\alpha)}}\Bigg[\sum_{L>1}\|w_N\|_{L_t^\infty \overline{H}^{-\frac12}}\|Q_{\sim LNN_1^\alpha} w_{N_1}\|_{F^{s-\frac{3-\alpha}{2},\frac12}}\|z_{N_1}\|_{L_t^2 H^s}\\ + \|w_N\|_{L_t^\infty \overline{H}^{-\frac12}}\|w_{N_1}\|_{L_t^2 H^{s-\frac{3-\alpha}{2}}}\|Q_{\sim NN_1^\alpha} z_{N_1}\|_{F^{s,\frac12}}\\ + N^{-\frac18}\langle N\rangle^{\frac{5+\alpha}{8}}N_1^{-\frac{\alpha}{8}-\frac12}\|w_N\|_{F^{-\frac12,\frac12}}\|w_{N_1}\|_{L_t^\infty H^{s-\frac{3-\alpha}{2}}}\|z_{N_1}\|_{L_t^2 H^s}\\ + N^{\frac14}N_1^{-\frac38}\|w_N\|_{L_t^\infty \overline{H}^{-\frac12}}\|w_{N_1}\|_{L_t^\infty H^{s-\frac{3-\alpha}{2}}}\|z_{N_1}\|_{L_t^\infty H^s}\Bigg]\\ \lesssim \|z\|_{Y^s}\big(\|w\|_{\widetilde{L}_t^\infty \overline{H}^{-\frac12}}\|w\|_{Z^{s-\frac{3-\alpha}{2}}} + \|w\|_{Z^{-\frac12}}\|w\|_{\widetilde{L}_t^\infty \overline{H}^{-\frac12}}\big),\end{aligned}$$

where again we used Cauchy–Schwarz in (N, N_1) and then summed over L.

Estimates for J_3. We first notice that for $N \lesssim N_1$ and $N_1 \gg 1$, since $1+2\big(s-\frac{3-\alpha}{2}\big) \geq 0$,

$$N\langle N^{-1}\rangle\langle N\rangle^{2(s-\frac{3-\alpha}{2})} \lesssim N_1\langle N_1^{-1}\rangle\langle N_1\rangle^{2(s-\frac{3-\alpha}{2})}.$$

Therefore, the contribution of this region to J_3 is controlled by J_2. Finally the contribution of $N \lesssim N_1 \lesssim 1$ is easily bounded by $\|z\|_{L^\infty_t L^2_x}\|w\|^2_{L^\infty_t \overline{H}^{-1/2}}$.

Gathering all the estimates, we eventually obtain

$$J \lesssim \|z\|_{Y^s}\|w\|_{L^\infty_T \overline{H}^{-\frac{1}{2}}}\|w\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|z\|_{Y_T^{1-\frac{\alpha}{2}}}\|w\|_{\widetilde{L}^\infty_T H^{s-\frac{3}{2}+\frac{\alpha}{2}}}\|w\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}, \tag{4-25}$$

which completes the proof of (4-22). □

4B. *Unconditional well-posedness.* Let us fix $s \geq 1-\frac{\alpha}{2}$. We notice that $1-\frac{\alpha}{2} \geq 0 > s_c = \frac{1}{2}-\alpha$, which is the critical Sobolev exponent associated with (1-3) for dilation symmetry. Therefore, as in Section 3B, the unconditional well-posedness in $H^s(\mathbb{R})$ of (1-3) for small H^s-initial data with a maximal time of existence $T \geq 1$ will ensure the unconditional well-posedness of (1-3) for arbitrary large H^s-initial data with a maximal time of existence

$$T \geq (1+\|u_0\|_{H^s})^{-\frac{2(\alpha+1)}{2\alpha-1}}.$$

Moreover, as in Section 3B, estimates (4-2), (4-3), (4-14), and a continuity argument ensure that a smooth solution with small H^s-initial datum has got a time of existence T in $H^\infty(\mathbb{R})$ that is greater than 1. Now, to prove the existence of a solution with initial data $u_0 \in H^{1-\alpha/2}$, we cannot argue exactly as in Section 3B since, for $s=0$, we miss compactness to pass to the limit on the nonlinear term. Instead, we construct below a sequence of smooth solutions to (1-3) that converges strongly to a solution of (1-3) emanating from u_0. This will be done by using the Bona–Smith argument.

Let $u_0 \in H^s$ with $s \geq 1-\frac{\alpha}{2}$ and $\|u_0\|_{H^s} \ll 1$. We denote by u^N the solution of (1-3) emanating from $P_{\leq N}u_0$. From the discussion above, $u_N \in C([0,1]; H^\infty(\mathbb{R}))$ and, for $1 \leq N_1 \leq N_2$, we set

$$w := u^{N_1} - u^{N_2}.$$

Let us note that $P_{\leq 1}w_0 = P_{\leq 1}(u^{N_1}-u^{N_2}) = 0$ and thus $w_0 \in \overline{H}^s(\mathbb{R})$ with $\|w_0\|_{\overline{H}^s} \sim \|w_0\|_{H^s}$. It then follows from (4-20)–(4-22) that

$$\|w\|_{Z_1^{s-\frac{3}{2}+\frac{\alpha}{2}}} \lesssim \|w(0)\|_{H^{s-\frac{3}{2}+\frac{\alpha}{2}}} \lesssim N_1^{\frac{\alpha-3}{2}}\|P_{>N_1}u_0\|_{H^s}. \tag{4-26}$$

Moreover, on account of Lemma 4.2, Proposition 4.5 and (4-19), for any $r \geq 0$ we have

$$\|u^{N_i}\|_{Y_T^{s+r}} \lesssim \|u^{N_i}\|_{\widetilde{Y}_T^{s+r}} \lesssim \|u_0^{N_i}\|_{H^{s+r}} \lesssim N_i^r\|u_0\|_{H^s}. \tag{4-27}$$

Next, since w satisfies (3-36), the Duhamel formula leads, for any $0 < N < 1$, to

$$\|P_N w\|_{L^\infty_1 \overline{H}^s} \lesssim \|P_N w_0\|_{\overline{H}^s} + N^{\frac{1}{2}}(\|u^{N_1}\|^2_{L^\infty_1 L^2_x} + \|w\|^2_{L^\infty_1 L^2_x})$$

and thus

$$\|P_{\leq 1}w\|_{\widetilde{L}^\infty_1 \overline{H}^s} \lesssim \|w_0\|_{H^s} + (\|u^{N_1}\|^2_{L^\infty_1 L^2_x} + \|w\|^2_{L^\infty_1 L^2_x}). \tag{4-28}$$

This proves that $w \in Z^s_T$. We will also need the following estimates on w:

Proposition 4.9. *Let $1 < \alpha \le 2$, $0 < T < 2$ and $w \in Z^s_T$ with $s \ge 1 - \frac{\alpha}{2}$ be a solution to (3-36). Then*

$$\|w\|_{Y^s_T} \lesssim \|w\|_{L^\infty_T H^s}(1 + \|u^{N_1}\|^2_{L^\infty_T H^s} + \|w\|^2_{L^\infty_T H^s}) \tag{4-29}$$

and

$$\|w\|^2_{\tilde{L}^\infty_T H^s} \lesssim \|w_0\|^2_{H^s} + \|w\|^3_{Y^s_T} + \|u^{N_1}\|_{Y^s_T}\|w\|^2_{Z^s_T} + \|u^{N_1}\|_{Y_T^{s+\frac{3}{2}-\frac{\alpha}{2}}}\|w\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}\|w\|_{Z^s_T}. \tag{4-30}$$

Proof. First, (4-29) can be derived exactly as (4-2)–(4-3) of Lemma 4.2. Now, to prove (4-30), we separate the contribution of $\partial_x(w^2)$ and $\partial_x(u^{N_1}w)$. First, (4-17) leads to

$$\sum_{N>0} N^{2s}\left|\int_0^t\int_{\mathbb{R}} P_N\partial_x(w^2)P_N w\right| \lesssim \|w\|^3_{Y^s_T}.$$

Second, applying (4-25) at the level s with z replaced by u^{N_1}, we obtain

$$\sum_{N>0} N^{2s}\left|\int_0^t\int_{\mathbb{R}} P_N\partial_x(u^{N_1}w)P_N w\right| \lesssim \|u^{N_1}\|_{Y_T^{s+\frac{3}{2}-\frac{\alpha}{2}}}\|w\|_{Z_T^{-\frac{1}{2}}}\|w\|_{Z^s_T} + \|u^{N_1}\|_{Y_T^{1-\frac{\alpha}{2}}}\|w\|^2_{Z^s_T},$$

which leads to (4-30) since $s - \frac{3}{2} + \frac{\alpha}{2} \ge -\frac{1}{2}$ for $s \ge 1 - \frac{\alpha}{2}$ and $Z^s_T \hookrightarrow Y^s_T$. $\square$

Combining (4-28), (4-29), (4-30) and (4-19), we infer that

$$\|w\|^2_{Z^s_1} \lesssim (1 + \|u_0\|^2_{H^s})^2\big[\|w_0\|^2_{H^s} + \|u_0\|_{H^s}\|w\|^2_{Y^s_1} + \|u_0\|_{H^s}\|w\|^2_{Z^s_1} + N_1^{\frac{3-\alpha}{2}}\|u_0\|_{H^s}\|w\|_{Z_1^{s-\frac{3}{2}+\frac{\alpha}{2}}}\|w\|_{Z^s_1}\big].$$

Then, the smallness assumption on $\|u_0\|_{H^s}$, (4-26) and the continuous injection $Z^s_T \hookrightarrow Y^s_T$, lead to

$$\begin{aligned}\|w\|^2_{Z^s_1} &\lesssim \|w_0\|^2_{H^s} + N_1^{3-\alpha}\|w\|^2_{Z_1^{s-\frac{3}{2}+\frac{\alpha}{2}}} \\ &\lesssim \|P_{>N_1}u_0\|^2_{H^s}(1 + \|P_{>N_1}u_0\|^2_{H^s}) \to 0 \quad \text{as } N_1 \to 0.\end{aligned} \tag{4-31}$$

This shows that $\{u^N\}$ is a Cauchy sequence in $C([0,1];H^s)$ and thus $\{u^N\}$ converges in $C([0,1];H^s)$ to a solution of (1-3) emanating from u_0. Note that there is no problem passing to the limit on the nonlinear term here, since we have strong convergence.

Now, Lemma 4.2, Proposition 4.5 and (4-19) ensure that any $L^\infty_1 H^s$-solution to (1-3) on $]0,1[$ belongs to $\tilde{Y}^s_T$. Therefore, according to Propositions 4.7 and 4.8, u is the only solution to (1-3) associated with the initial datum u_0 that belongs to $L^\infty_{\mathrm{loc}} H^s$.

To prove the continuity of the solution map in $H^s(\mathbb{R})$, we proceed as in Section 3B. Let $\{u_{0,n}\} \subset H^s(\mathbb{R})$ be such that $u_{0,n} \to u_0$ in $H^s(\mathbb{R})$ and let $\{u_n\} \subset C([0,1];H^s(\mathbb{R}))$ be the associated sequence of solutions to (1-3). Taking the same notations as above, we observe that, by construction,

$$P_{\le 1}(u_0 - u_0^N) = P_{\le 1}(u_{0,n} - u_{0,n}^N) = 0 \quad \text{for all } N \ge 1.$$

This ensures that $u-u^N$ and $u_n-u_n^N$ belong to Z^s_T. Estimate (4-31) on solutions to (3-36) then leads to

$$\|u-u^N\|_{Z^s_1}+\|u_n-u_n^N\|_{Z^s_1}\lesssim \|P_{>N}u_0\|_{H^s}+\|P_{>N}u_{0,n}\|_{H^s},$$

which yields

$$\lim_{N\to\infty}\sup_{n\in\mathbb{N}}(\|u-u^N\|_{L^\infty_1H^s}+\|u_n-u_n^N\|_{L^\infty_1H^s})=0. \tag{4-32}$$

It remains to estimate $\|u_n^N-u^N\|_{H^s}$. Note that we cannot use Propositions 4.8 and 4.9 here, since $u_{0,n}^N-u_0^N$ does not belong a priori to $\overline{H}^s(\mathbb{R})$. However, since u_0^N and $u_{0,n}^N$ belong to $H^\infty(\mathbb{R})$, we know, from the beginning of this section, that u^N and u_n^N belong to $C([0,1];H^\infty(\mathbb{R}))$. We now fix $N\gg 1$. Setting $s'=\max(1,s)$, we have

$$\|u_0^N-u_{0,n}^N\|_{H^{s'}}\to 0 \quad \text{as } n\to\infty.$$

Therefore, on account of Section 3B,

$$\|u^N-u_n^N\|_{L^\infty_TH^{s'}}\to 0 \quad \text{as } n\to\infty \quad \text{with} \quad T\sim(1+\|u_0^N\|_{H^{s'}})^{-\frac{2(\alpha+1)}{2\alpha-1}}.$$

Since $u^N\in C([0,1];H^\infty(\mathbb{R}))$ we can iterate this argument a finite number of times to obtain that the convergence of u_n^N to u^N holds actually in $C([0,1];H^{s'}(\mathbb{R}))$. The continuity of the flow map in $H^s(\mathbb{R})$ follows by combining this last result with (4-32).

4C. ***The periodic case.*** We use the notations of Section 3C. Let $H^s_0(\lambda\mathbb{T})$ be the closed subspace of zero-mean functions of $H^s_0(\lambda\mathbb{T})$. We define the Banach space $\overline{H}^s(\lambda\mathbb{T})$ as the space $H^s_0(\lambda\mathbb{T})$ endowed with the norm

$$\|u\|_{\overline{H}^s}=\|\langle|\xi|^{-\frac12}\rangle\langle\xi\rangle^s\hat{\varphi}\|_{L^2_\xi}.$$

Let $(u,v)\in(L^\infty(0,T;H^s(\lambda\mathbb{T})))^2$ be a pair of solutions to (1-3) associated with initial data (u_0,v_0) in $(H^s(\lambda\mathbb{T}))^2$ such that $u_0-v_0\in\overline{H}^s(\lambda\mathbb{T})$. As noticed in Remark 1.3, $(u,v)\in C([0,T];H^{s-\alpha-1}(\lambda\mathbb{T}))^2$ and it is not too hard to check that the mean value is a constant of the motion for such solutions. Therefore, $u(t)-v(t)$ has mean value zero for all $t\in[0,T]$.

As explained in Section 3C, to extend our result on the torus $\lambda\mathbb{T}$, uniformly for $\lambda\geq 1$, we only have to care about the contributions of the null frequencies each time we used the homogeneous decomposition in space frequencies. First we notice that in the proof of Lemma 4.2 we do not use any homogeneous decomposition in space frequencies and thus this lemma still holds in the periodic setting. Note that this is also true for (4-29), since the proof of this estimate is exactly the same. Moreover, on account of (3-42), the contributions of the null frequencies vanish in the proof of Lemma 4.2. Now, for Propositions 4.7, 4.8 and 4.9, we only have to care about the contributions of $\partial_x P_N(wP_0z)$, since, according to the discussion above, $P_0w=P_0(u-v)=0$ on $[0,T]$. On account of (3-43), these contributions vanish in (4-22) and (4-30). Finally, these contributions can be estimated in Proposition 4.7 by

$$\|\partial_xP_N(P_NwP_0z)\|_{F^{s-\frac32+\frac{\alpha}{2},-\frac12}}\lesssim N\|P_N(P_NwP_0z)\|_{X^{s-2,0}}\lesssim\delta_N\|z\|_{L^\infty_tL^2_x}\|w\|_{L^2_tH^{s-1}_x}$$

with $\|(\delta_{2^j})\|_{l^1(\mathbb{Z})} \lesssim 1$. This is acceptable, since $1-\frac{\alpha}{2} \geq 0$ and $s-\frac{3}{2}+\frac{\alpha}{2} \geq s-1$. The proof of Theorem 1.5 is now complete.

5. Dissipative limits

First, we notice that, if u is a solution to (1-9), then u_λ defined by $u_\lambda(t,x) = \lambda^\alpha u(\lambda^{1+\alpha}t, \lambda x)$ is a solution to

$$\partial_t u_\lambda + L^\lambda_{\alpha+1} u_\lambda + \varepsilon \lambda^{\alpha+1-\beta} A^\lambda_\beta u_\lambda + \tfrac{1}{2}\partial_x (u_\lambda)^2 = 0 \tag{5-1}$$

with

$$\widehat{L^\lambda_{\alpha+1} v}(\xi) = i\lambda^{\alpha+1} p_{\alpha+1}(\lambda^{-1}\xi)\hat{v}(\xi)$$

and

$$\widehat{A^\lambda_\beta v}(\xi) = \lambda^\beta q_\beta(\lambda^{-1}\xi)\hat{v}(\xi) \quad \text{for all } \xi \in \mathbb{R}.$$

Therefore, as in the preceding section, up to this change of unknown, of parameter ε and of operators, we may assume that u satisfies (1-9) with $L_{\alpha+1}$ and A_β that verify Hypotheses 1 and 2 for all $0 < \lambda \leq 1$.

Second, we notice that Hypothesis 2 now ensures that, for $0 < \lambda \leq 1$ and $N \gg 1$ dyadic,

$$(A^\lambda_\beta P_N v, P_N v)_{L^2} \gtrsim N^{\frac{\beta}{2}} \|P_N v\|^2_{L^2} \tag{5-2}$$

and

$$\|A^\lambda_\beta P_N v\|_{L^2} \lesssim N^\beta \|P_N v\|_{L^2}. \tag{5-3}$$

The main point is now to prove that the Cauchy problem (1-9) is locally well-posed in H^s uniformly in $\varepsilon > 0$.

Proposition 5.1. *Let $1 \leq \alpha \leq 2$, $0 \leq \beta \leq 1+\alpha$ and $s \geq 1-\frac{\alpha}{2}$. For any $\varphi \in H^s(\mathbb{R})$ there exists $T \sim (1+\|u_0\|_{H^{1-\alpha/2}})^{-2(\alpha+1)/(2\alpha-1)}$ and a solution $u_\varepsilon \in C([0,T]; H^s)$ to (1-9) that is unique in some function space[2] embedded in $L^\infty_T(0,T; H^s)$. Moreover, there exists $C > 0$ such that, for any $\varepsilon \in]0,1[$,*

$$\sup_{t\in[0,T]} \|u_\varepsilon(t)\|_{H^s} \leq C\|\varphi\|_{H^s}. \tag{5-4}$$

Finally, for any $R > 0$, the family of solution maps $S_\varepsilon : \varphi \mapsto u_\varepsilon$, $\varepsilon \in]0,1[$, from $B(0,R)_{H^s}$ into $C([0,T(R)]; H^s(\mathbb{R}))$ is equicontinuous, i.e., for any sequence $\{\varphi_n\} \subset B(0,R)_{H^s}$ converging to φ in $H^s(\mathbb{R})$,

$$\lim_{n\to+\infty} \sup_{\varepsilon\in]0,1[} \|S_\varepsilon\varphi - S_\varepsilon\varphi_n\|_{L^\infty(0,T(R);H^s(\mathbb{R}))} = 0. \tag{5-5}$$

Proof. We treat the cases $(\alpha,s) \neq \left(1,\frac{1}{2}\right)$. This last case can be treated in the same way by using the estimates derived in the Appendix. First we notice that, for (1-9), in view of (5-2), the energy estimate (4-14) becomes

$$\|u\|_{\tilde{L}^\infty_T H^s} + \sqrt{\varepsilon}\|u\|_{L^2_T H^{s+\frac{\beta}{2}}} \lesssim \|u_0\|_{H^s} + \|u\|_{L^\infty_T H^{1-\frac{\alpha}{2}}} \|u\|_{Y^s_T} + \|u\|_{L^\infty_T H^s} \|u\|_{Y^{1-\frac{\alpha}{2}}_T}. \tag{5-6}$$

[2]For $(\alpha,s) \neq \left(1,\frac{1}{2}\right)$, this space is simply the space $L^\infty_T H^s \cap L^2_T H^{s+\beta/2}$.

On the other hand, viewing $\varepsilon A_\beta u$ as a forcing term, (4-2)–(4-3) together with (5-3) lead to

$$\|u\|_{Y^s_T} \lesssim \|u\|_{L^\infty_T H^s}(1+\|u\|^2_{L^\infty_T H^{1-\frac{\alpha}{2}}}) + \varepsilon\|u\|_{L^2_T H^{s-\frac{1+\alpha}{2}+\beta}}. \tag{5-7}$$

To derive an a priori bound from the above estimates, as in the previous section, we have to use the dilation argument that is described in the beginning of this section. So the dilation function u_λ defined by $u_\lambda(t,x)=\lambda^\alpha u(\lambda^{1+\alpha}t,\lambda x)$ satisfies (5-1) and we set

$$\|v\|_{N^s} := \|v\|_{L^\infty_T H^s} + \sqrt{\varepsilon\lambda^{\alpha+1-\beta}}\,\|v\|_{L^2_T H^{s+\frac{\beta}{2}}}.$$

Since $\beta \le \alpha+1$, this ensures that, for $\lambda \lesssim (1+\|\varphi\|_{H^s})^{-2(\alpha+1)/(2\alpha-1)}$ and $0<T\le 2$,

$$\|u_\lambda\|_{N^s_T} \lesssim \|\varphi_\lambda\|_{H^s} + (1+\|u_\lambda\|^2_{N_T^{1-\frac{\alpha}{2}}})\|u_\lambda\|_{N_T^{1-\frac{\alpha}{2}}}\|u_\lambda\|_{N^s_T}$$

with $\|\varphi_\lambda\|_{H^s} \lesssim \lambda^{\alpha-1/2}\|\varphi\|_{H^s} \ll 1$. This leads to the uniform bound (5-4) for smooth solutions to (1-9) by a classical continuity argument.

Now, proceeding in the same way for the difference of two solutions, it is not too hard to check that (4-20) becomes

$$\begin{aligned}&\|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}\\ &\lesssim \|u-v\|_{\widetilde{L}^\infty_T \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|u-v\|_{L^2_T H^{s-\frac{3}{2}+\frac{\alpha}{2}+\beta}} + \|u+v\|_{\widetilde{Y}^s_T}\|u-v\|_{Z_T^{-\frac{1}{2}}} + \|u+v\|_{\widetilde{Y}^{1-\frac{\alpha}{2}}}\|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}},\end{aligned}$$

whereas (4-22) becomes

$$\|u-v\|_{\widetilde{L}^\infty_T \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \sqrt{\varepsilon}\|u-v\|_{L^2_T \overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}+\frac{\beta}{2}}} \lesssim \|u_0-v_0\|_{\overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \|u+v\|_{\widetilde{Y}^s_T}\|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}}.$$

By the same dilation arguments as above, this leads to

$$\|u-v\|_{Z_T^{s-\frac{3}{2}+\frac{\alpha}{2}}} + \sqrt{\varepsilon}\|u-v\|_{L^2_T H^{s-\frac{3}{2}+\frac{\alpha}{2}+\frac{\beta}{2}}} \lesssim \|u_0-v_0\|_{\overline{H}^{s-\frac{3}{2}+\frac{\alpha}{2}}}. \tag{5-8}$$

Combining the above estimates and the Bona–Smith argument, we can proceed as in Section 4B and construct a sequence of smooth solutions that converges strongly in $C([0,T];H^s)$ towards a solution u_ε to (1-9). We thus obtain the existence of a solution $u_\varepsilon \in C([0,T];H^s)\cap L^2_T H^{s+\beta/2}$ to (1-9) with $T \gtrsim (1+\|u_0\|_{H^{1-\alpha/2}})^{-2(\alpha+1)/(2\alpha-1)}$ and $\varphi \in H^s$ as initial data. Moreover, (5-8) ensures that this is the only solution emanating from φ in the class $L^\infty_{\text{loc}} H^s \cap L^2_{\text{loc}} H^{s+\beta/2}$. Obviously, this solution satisfies (5-4). Finally, the equicontinuity of the solution map in $C(0,T;H^s)$ follows from Bona–Smith arguments as in Section 3B. □

It is clear that the above proposition implies part (1) of Theorem 1.14. Now, part (2) will follow from general arguments (see for instance [Guo and Wang 2009]). Let us denote by S_ε and S the nonlinear group associated with, respectively, (1-9) and (1-3). Let $\varphi \in H^s(\mathbb{R})$, $s \ge 1-\frac{\alpha}{2}$ and let $T = T(\|\varphi\|_{H^{1-\alpha/2}}) > 0$

be as given by Proposition 5.1. For any $N>0$ we can rewrite $S_\varepsilon(\varphi)-S(\varphi)$ as

$$\begin{aligned} S_\varepsilon(\varphi)-S(\varphi) &= (S_\varepsilon(\varphi)-S_\varepsilon(P_{\le N}\varphi))+(S_\varepsilon(P_{\le N}\varphi)-S(P_{\le N}\varphi))+(S(P_{\le N}\varphi)-S(\varphi)) \\ &= I_{\varepsilon,N}+J_{\varepsilon,N}+K_N. \end{aligned}$$

By continuity with respect to initial data in $H^s(\mathbb{R})$ of the solution map associated with (1-3), we have $\lim_{N\to\infty}\|K_N\|_{L^\infty(0,T;H^s)}=0$. Moreover, (5-5) ensures that

$$\lim_{N\to\infty}\sup_{\varepsilon\in]0,1[}\|I_{\varepsilon,N}\|_{L^\infty(0,T;H^s)}=0.$$

It thus remains to check that, for any fixed $N>0$, $\lim_{\varepsilon\to 0}\|J_{\varepsilon,N}\|_{L^\infty(0,T;H^s_x)}=0$. Since $P_{\le N}\varphi\in H^\infty(\mathbb{R})$, it is worth noticing that $S_\varepsilon(P_{\le N}\varphi)$ and $S(P_{\le N}\varphi)$ belong to $C^\infty(\mathbb{R};H^\infty(\mathbb{R}))$. Moreover, according to Theorem 1.14 and Proposition 5.1, for all $\theta\in\mathbb{R}$ and $\varepsilon\in]0,1[$,

$$\|S_\varepsilon(P_{\le N}\varphi)\|_{L^\infty_T H^\theta_x}+\|S(P_{\le N}\varphi)\|_{L^\infty_T H^\theta_x}\le C(N,\theta,\|\varphi\|_{L^2_x}).$$

Now, setting $v_\varepsilon:=S_\varepsilon(P_{\le N}\varphi)$ and $v:=S(P_{\le N}\varphi)$, we observe that $w_\varepsilon:=v_\varepsilon-v$ satisfies

$$\partial_t w_\varepsilon+L_{\alpha+1}w_\varepsilon=-\tfrac{1}{2}\partial_x(w_\varepsilon(v+v_\varepsilon))-\varepsilon A_\beta v_\varepsilon$$

with initial data $w_\varepsilon(0)=0$. For $s\ge 0$, taking the H^s-scalar product of this last equation with w_ε and integrating by parts, we get

$$\frac{d}{dt}\|w_\varepsilon\|_{H^s}\lesssim(1+\|\partial_x(v+v_\varepsilon)\|_{L^\infty_x})\|w_\varepsilon\|^2_{H^s}+\|[J^s\partial_x,(v+v_\varepsilon)]w_\varepsilon\|_{L^2_x}\|w_\varepsilon\|_{H^s}+\varepsilon^2\|D^\beta_x v_\varepsilon\|^2_{H^s}.$$

Applying the mean value theorem to the Fourier transform of the commutator term, it is not too hard to check that

$$\|[J^s_x\partial_x,f]g\|_{L^2_x}\lesssim\|f_x\|_{H^{s+1}}\|g\|_{H^s_x}, \tag{5-9}$$

which leads to

$$\frac{d}{dt}\|w_\varepsilon(t)\|^2_{H^s}\lesssim C(N,s+2,\|\varphi\|_{L^2_x})\|w_\varepsilon(t)\|^2_{H^s_x}+\varepsilon^2 C(N,s+\beta,\|\varphi\|_{L^2_x})^2.$$

Integrating this differential inequality on $[0,T]$, this ensures that $\lim_{\varepsilon\to 0}\|w_\varepsilon\|_{L^\infty(0,T;H^s)}=0$ and proves that

$$u_\varepsilon\to u\quad\text{in } C([0,T];H^s) \tag{5-10}$$

with $T\sim(1+\|u_0\|_{H^{1-\alpha/2}})^{-2(\alpha+1)/(2\alpha-1)}$. Now fix $\varphi\in H^s$ and let $T^*>0$ be the maximal time of existence of $S(\varphi)$. It remains to prove that the time of existence T_ε of $S_\varepsilon(\varphi)$ in H^s satisfies $\liminf_{\varepsilon\to 0}T_\varepsilon\ge T^*$. Actually, this follows by a classical contradiction argument. Indeed, assuming that this is not true, there exist $\varepsilon_n\searrow 0$ such that $\lim T_{\varepsilon_n}=T_1<T^*$. We set

$$\delta(T_1)=(1+\|S(\varphi)\|_{L^\infty(0,T_1;H^{1-\frac{\alpha}{2}})})^{-\frac{2(\alpha+1)}{2\alpha-1}},$$

which is well defined since $T_1 < T^*$. Applying (5-10) about $T_1/\delta(T_1)$ times, we eventually obtain that, for n large enough,

$$\big\| S_{\varepsilon_n}(\varphi)\big(T_1 - \tfrac{1}{100}\delta(T_1)\big)\big\|_{H^{1-\frac{\alpha}{2}}} \le 2\|S(\varphi)\|_{L^\infty(0,T_1;H^{1-\frac{\alpha}{2}})}.$$

But then the uniform bound from below on the existence time ensures that $T_{\varepsilon_n} \ge T_1 + \frac{1}{2}\delta(T_1)$, which contradicts $\lim T_{\varepsilon_n} = T_1$ and proves the desired result. This ensures that, fixing $0 < T_0 < T^*$, we have $T_\varepsilon \ge T_0$ for $\varepsilon > 0$ small enough. Finally, applying (5-10) about $T_0/\delta(T_0)$ times, we get (5-10) with $T = T_0$. This completes the proof of Theorem 1.14.

Appendix: The case $\alpha = 1$ and $s = \frac{1}{2}$

This case is important since $H^{1/2}$ is the energy space for the Benjamin–Ono equation and also the intermediate long waves equation. Unfortunately, we are not able to prove the unconditional well-posedness in this case. However, we are able to prove the well-posedness without using a gauge transform. This is useful for treating perturbations of these equations, as we explained in the preceding section. In this section, we indicate the modifications of the proofs in this case. In the sequel we set

$$\widetilde{M}^{\frac{1}{2}} := \widetilde{L}^\infty_t H^{\frac{1}{2}} \cap X^{-\frac{1}{2},1}.$$

Lemma A.1. *Let $\alpha = 1$, $0 < T < 2$, and let $u \in \widetilde{M}^{1/2}_T$ be a solution to (1-3). Then*

$$\|u\|_{\widetilde{M}^{\frac{1}{2}}_T} \lesssim \|u\|_{\widetilde{L}^\infty_T H^{\frac{1}{2}}} + \|u\|^2_{\widetilde{M}^{\frac{1}{2}}_T}. \tag{A-1}$$

Proof. Working with the extension $\tilde{u} = \rho_T u$ (see (3-3)), still denoted u, if suffices to estimate the $X^{-1/2,1}$-norm of u. First we notice that the low frequency part can be easily controlled by

$$\|P_{\lesssim 1} u\|_{X_T^{-\frac{1}{2},1}} \lesssim \|u\|^2_{L^\infty_T L^2_x}.$$

Now, for $N \gg 1$, we have

$$\begin{aligned}\|u_N\|_{X_T^{-\frac{1}{2},1}} &\lesssim \|P_N u_0\|_{H^{-\frac{1}{2}}} + N^{\frac{1}{2}} \Big\| \sum_{N_2' \sim N_2 \gtrsim N} u_{N_2} u_{N_2'} \Big\|_{L^2_T L^2_x} \\ &\qquad + N^{\frac{1}{2}} \Big\| \sum_{1 \le N_2 \ll N} P_N(u_{\sim N} u_{N_2}) \Big\|_{L^2_T L^2_x} + N^{\frac{1}{2}} \Big\| \sum_{N_2 < 1} P_N(u_{\sim N} u_{N_2}) \Big\|_{L^2_T L^2_x} \\ &= \|P_N u_0\|_{H^{-\frac{1}{2}}} + I_N + II_N + III_N.\end{aligned}$$

Clearly,

$$\begin{aligned} I_N &\lesssim N^{\frac{1}{2}} \sum_{N_2' \sim N_2 \gtrsim N} \|u_{N_2}\|_{L^2_{tx}} \|u_{N_2'}\|_{L^\infty_t H^{\frac{1}{2}}} \\ &\lesssim \|u\|_{L^\infty_t H^{\frac{1}{2}}} \sum_{N_2 \gtrsim N} \Big(\frac{N}{N_2}\Big)^{\frac{1}{2}} \|u_{N_2}\|_{L^2_t H^{\frac{1}{2}}} \\ &\lesssim \delta_N \|u\|^2_{L^\infty_t H^{\frac{1}{2}}} \end{aligned}$$

with $\|(\delta_{2^j})\|_{l^2(\mathbb{N}^*)} \lesssim 1$. Moreover, we easily get from Bernstein estimates that

$$III_N \lesssim N^{\frac{1}{2}} \sum_{N_2<1} \|u_{\sim N}\|_{L^2_{tx}} \|u_{N_2}\|_{L^\infty_{tx}} \lesssim \|u_{\sim N}\|_{L^2_t H^{\frac{1}{2}}} \|u\|_{L^\infty_t H^{\frac{1}{2}}} \lesssim \delta_N \|u\|_{L^\infty_t H^{\frac{1}{2}}} \|u\|_{L^\infty_t H^{\frac{1}{2}}}$$

with $\|(\delta_{2^j})\|_{l^2(\mathbb{N}^*)} \lesssim 1$. On the other hand,

$$II_N \lesssim N^{\frac{1}{2}} \Bigl\| \sum_{1\le N_2\ll N} Q_{\sim NN_2} P_N(u_{\sim N} u_{N_2}) \Bigr\|_{L^2_{tx}} + N^{\frac{1}{2}} \Bigl\| \sum_{1\le N_2\ll N} Q_{\nsim NN_2} P_N(u_{\sim N} u_{N_2}) \Bigr\|_{L^2_{tx}} \lesssim II^1_N + II^2_N.$$

By almost orthogonality, we have

$$\begin{aligned}
II^1_N &\lesssim N^{\frac{1}{2}} \Bigl(\sum_{1\le N_2\ll N} \bigl\| Q_{\sim NN_2} P_N(u_{\sim N} u_{N_2}) \bigr\|^2_{L^2_{tx}} \Bigr)^{\frac{1}{2}} \\
&\lesssim N^{\frac{1}{2}} \Bigl(\sum_{N_2\ll N} \|u_{\sim N}\|^2_{L^2_{tx}} \|u_{N_2}\|^2_{L^\infty_t H^{\frac{1}{2}}_x} \Bigr)^{\frac{1}{2}} \\
&\lesssim \|u_{\sim N}\|_{L^2_t H^{\frac{1}{2}}} \|u\|_{\widetilde{L}^\infty_t H^{\frac{1}{2}}} \\
&\lesssim \delta_N \|u\|_{L^\infty_t H^{\frac{1}{2}}} \|u\|_{\widetilde{L}^\infty_t H^{\frac{1}{2}}}
\end{aligned}$$

with $\|(\delta_{2^j})\|_{l^2(\mathbb{N}^*)} \lesssim 1$. It remains to control II^2_N. Since the Fourier projectors ensure $\langle \tau - p_2(\xi)\rangle \nsim NN_2$, the resonance relation (1-6) leads to $|\tau_1 - p_2(\xi_1)| \vee |\tau - \tau_1 - p_2(\xi-\xi_1)|) \gtrsim NN_2$ for II^2_N. We separate the contributions of $Q_{\gtrsim NN_2} u_{\sim N}$ and $Q_{\gtrsim NN_2} u_{N_2}$. For the first contribution, we have

$$\begin{aligned}
II^2_N &\lesssim N^{\frac{1}{2}} \sum_{1\le N_2\ll N} (NN_2)^{-\frac{1}{4}} N^{\frac{1}{4}} \|Q_{\gtrsim NN_2} u_{\sim N}\|_{X^{\frac{1}{4},\frac{1}{4}}} \|u_{N_2}\|_{L^\infty_t H^{\frac{1}{2}}} \\
&\lesssim \|u_{\sim N}\|_{X^{\frac{1}{4},\frac{1}{4}}} \|u\|_{L^\infty_t H^{\frac{1}{2}}} \\
&\lesssim \delta_N \|u\|^{\frac{1}{4}}_{X^{-\frac{1}{2},1}} \|u\|^{\frac{3}{4}}_{L^\infty_t H^{\frac{1}{2}}} \|u\|_{L^\infty_t H^{\frac{1}{2}}}
\end{aligned}$$

with $\|(\delta_{2^j})\|_{l^2(\mathbb{N}^*)} \lesssim 1$ and where we used interpolation at the last step. For the second contribution, we write

$$\begin{aligned}
II^2_N &\lesssim N^{\frac{1}{2}} \sum_{1\le N_2\ll N} \|Q_{\ll NN_2} u_{\sim N}\|_{L^\infty_t L^4_x} \|Q_{\gtrsim NN_2} u_{N_2}\|_{L^2_t L^4_x} \\
&\lesssim N^{\frac{1}{2}} \sum_{1\le N_2\ll N} N^{-\frac{1}{4}} \|Q_{\ll NN_2} u_{\sim N}\|_{L^\infty_t H^{\frac{1}{2}}} \|Q_{\gtrsim NN_2} u_{N_2}\|_{L^2_t H^{\frac{1}{4}}} \\
&\lesssim N^{\frac{1}{2}} \sum_{1\le N_2\ll N} N^{-\frac{1}{4}} (NN_2)^{-\frac{1}{4}} \|u_{\sim N}\|_{L^\infty_t H^{\frac{1}{2}}} \|u_{N_2}\|_{X^{\frac{1}{4},\frac{1}{4}}} \\
&\lesssim \delta_N \|u\|_{\widetilde{L}^\infty_t H^{\frac{1}{2}}} \|u\|^{\frac{1}{4}}_{X^{-\frac{1}{2},1}} \|u\|^{\frac{3}{4}}_{L^\infty_t H^{\frac{1}{2}}}
\end{aligned}$$

with $\|(\delta_{2^j})\|_{l^2(\mathbb{N}^*)} \lesssim 1$. Gathering the above estimates, (5-2) follows. □

Lemma A.2. *Let $\alpha = 1$, $0 < T < 2$ and let $u \in \widetilde{M}_T^{1/2}$ be a solution to (1-3). Then*

$$\|u\|^2_{\widetilde{L}^\infty_T H^{\frac12}} \lesssim \|u_0\|^2_{H^{\frac12}} + \|u\|^2_{\widetilde{L}^\infty_t H^{\frac12}} \|u\|_{\widetilde{M}_T^{\frac12}}. \tag{A-2}$$

Proof. We follow the proof of Proposition 4.5. Note that $\widetilde{M}^{1/2} \hookrightarrow \widetilde{Y}^{1/2}$. According to (4-15), it suffices to control

$$I = \sum_{N>0} \sum_{N_1 \gtrsim N} N\langle N_1\rangle \sup_{t\in]0,T[} |I_t(u_N, u_{\sim N_1}, u_{N_1})|.$$

It is easy to check that the only term of the left-hand side of (4-16) that causes trouble in the case $\alpha = 1$ is the first one. This term corresponds to the contribution of $Q_{\sim LNN_1^\alpha} u_{N_1}$ and $Q_{\sim NN_1^\alpha} u_{\sim N_1}$. For $\alpha = 1$, we control these contributions by applying Cauchy–Schwarz in (N, N_1). For instance, the contribution of $Q_{\sim LNN_1^\alpha} u_{N_1}$ is estimated, thanks to Lemma 4.4, by

$$\begin{aligned}
&\sum_{N\gg 1} \sum_{N_1\gtrsim N} N\langle N_1\rangle \sum_{L>1} L^{-1} N^{-\frac12} \|u_N\|_{L^2_{tx}} \|Q_{\sim LNN_1^\alpha} u_{N_1}\|_{F^{0,\frac12}} \|u_{\sim N_1}\|_{L^\infty_t L^2_x} \\
&\qquad \lesssim \sum_{L>1} L^{-1} \Big(\sum_{N_1\gtrsim N\gg 1} \|u_N\|^2_{L^2_t H^{\frac12}} \|u_{\sim N_1}\|^2_{L^\infty_t H^{\frac12}} \Big)^{\frac12} \Big(\sum_{N_1\gtrsim N\gg 1} \|Q_{\sim LNN_1^\alpha} u_{N_1}\|^2_{F^{\frac12,\frac12}} \Big)^{\frac12} \\
&\qquad \lesssim \|u\|_{L^2_t H^{\frac12}} \|u\|_{\widetilde{L}^\infty_t H^{\frac12}} \|u\|_{X^{-\frac12,1}}. \qquad \square
\end{aligned}$$

Lemma A.3. *Let $0 < T < 2$ and let $u, v \in \widetilde{M}_T^{1/2}$ be two solutions to (1-3) on $]0, T[$. Then we have*

$$\|u - v\|_{Z_T^{-\frac12}} \lesssim \|u - v\|_{L^\infty_T \overline{H}^{s-\frac32+\frac{\alpha}{2}}} + \|u + v\|_{\widetilde{M}_T^{\frac12}} \|u - v\|_{Z_T^{-\frac12}} \tag{A-3}$$

and

$$\|u - v\|^2_{\widetilde{L}^\infty_T \overline{H}^{-\frac12}} \lesssim \|u_0 - v_0\|^2_{\overline{H}^{-\frac12}} + \|u + v\|_{\widetilde{M}_T^{\frac12}} \|u - v\|_{\widetilde{L}^\infty_T \overline{H}^{-\frac12}} \|u - v\|_{Z_T^{-\frac12}}. \tag{A-4}$$

Proof. First we notice that (A-4) is already proven in Proposition 4.8, since $\widetilde{M}_T^{1/2} \hookrightarrow \widetilde{Y}_T^{1/2} \hookrightarrow Y_T^{1/2}$. It remains to prove (A-3). We follow the proof of Proposition 4.5. It is not too hard to check that the only contribution that causes troubles in the right-hand side of (4-21), in the case $\alpha = 1$, is the contribution of the low-high interaction term, $P_N(P_{\lesssim N} z w_N)$. We proceed as in Lemma A.1 . We take extensions $\tilde{z}$ and $\widetilde{w}$, supported in $]-4, 4[$, of z and w such that $\|\tilde{z}\|_{\widetilde{M}^{1/2}} \lesssim \|z\|_{\widetilde{M}_T^{1/2}}$ and $\|\widetilde{w}\|_{Z^{-1/2}} \lesssim \|w\|_{Z_T^{-1/2}}$. For simplicity we drop the tilde. We first notice that the contribution of $P_{\lesssim 1} z$ is easily estimated by

$$\|\partial_x P_N(P_{\lesssim 1} z w_{\sim N})\|_{F^{-\frac12,1,-\frac12}} \lesssim \langle N\rangle^{-\frac12} \|P_N(P_{\lesssim 1} z w_{\sim N})\|_{L^2_{tx}} \lesssim \|z\|_{L^\infty_t L^2_x} \|w_{\sim N}\|_{L^2_t H^{-\frac12}},$$

which is acceptable. Now we decompose the remaining contribution as

$$
\begin{aligned}
&\|\partial_x P_N(P_{\gg 1} P_{\lesssim N} z w_{\sim N})\|_{F^{-\frac{1}{2},1,-\frac{1}{2}}} \\
&\quad \lesssim N \Big\| \sum_{1 \ll N_1 \lesssim N} P_N(P_{N_1} z w_{\sim N}) \Big\|_{X^{-\frac{3}{2},0}} \\
&\quad \lesssim \langle N \rangle^{-\frac{1}{2}} \Big\| \sum_{1 \ll N_1 \lesssim N} Q_{\sim N N_1} P_N(P_{N_1} z w_{\sim N}) \Big\|_{L^2_{tx}} + \langle N \rangle^{-\frac{1}{2}} \Big\| \sum_{1 \ll N_1 \lesssim N} Q_{\nsim N N_1} P_N(P_{N_1} z w_{\sim N}) \Big\|_{L^2_{tx}} \\
&\quad = J_{1,N} + J_{2,N}.
\end{aligned}
$$

By almost-orthogonality,

$$
\begin{aligned}
J_{1,N} &\lesssim \langle N \rangle^{-\frac{1}{2}} \Big(\sum_{1 \ll N_1 \lesssim N} \|Q_{\sim N N_1} P_N(P_{N_1} z w_{\sim N})\|^2_{L^2_{tx}} \Big)^{\frac{1}{2}} \\
&\lesssim \langle N \rangle^{-\frac{1}{2}} \Big(\sum_{1 \ll N_1 \lesssim N} \|P_{N_1} z\|^2_{L^2_t H^{\frac{1}{2}}} \|w_{\sim N}\|^2_{L^\infty_{tx}} \Big)^{\frac{1}{2}} \\
&\lesssim \|w_{\sim N}\|_{L^\infty_t H^{-\frac{1}{2}}} \|z\|_{L^2_t H^{\frac{1}{2}}},
\end{aligned}
$$

which is acceptable. To treat J_2, we notice that, since the Fourier projectors ensure that $\langle \tau - p_2(\xi) \rangle \nsim N N_1$, the resonance relation (1-6) leads to $|\tau_1 - p_2(\xi_1)| \vee |\tau - \tau_1 - p_2(\xi - \xi_1)| \gtrsim N N_1$ for $J_{2,N}$. We separate the contributions of $Q_{\gtrsim N N_1} z_{N_1}$ and $Q_{\gtrsim N N_1} w_{\sim N}$. For the first contribution, we write

$$
\begin{aligned}
J_{2,N} &\lesssim \langle N \rangle^{-\frac{1}{2}} \sum_{1 \ll N_1 \lesssim N} N_1^{\frac{1}{2}} \|Q_{\gtrsim N N_1} P_{N_1} z\|_{L^2_{tx}} \|w_{\sim N}\|_{L^\infty_t L^2_x} \\
&\lesssim \langle N \rangle^{-\frac{1}{2}} \sum_{1 \ll N_1 \lesssim N} (N N_1)^{-\frac{1}{4}} N_1^{\frac{1}{4}} \|Q_{\gtrsim N N_1} P_{N_1} z\|_{X^{\frac{1}{4},\frac{1}{4}}} \|w_{\sim N}\|_{L^\infty_t L^2_x} \\
&\lesssim \|z\|^{\frac{1}{4}}_{X^{-\frac{1}{2},1}} \|z\|^{\frac{3}{4}}_{L^\infty_t H^{\frac{1}{2}}} \|w_{\sim N}\|_{L^\infty_t H^{-\frac{1}{2}}},
\end{aligned}
$$

which is acceptable. For the second contribution, according to (4-10), we have

$$
\begin{aligned}
J_2 &\lesssim \langle N \rangle^{-\frac{1}{2}} \sum_{1 \ll N_1 \lesssim N} \|z_{N_1}\|_{L^\infty_t H^{\frac{1}{2}}} \|Q_{\gtrsim N N_1} w_{\sim N}\|_{L^2_{tx}} \\
&\lesssim \langle N \rangle^{-\frac{1}{2}} \sum_{1 \ll N_1 \lesssim N} (N N_1)^{-1} N^{\frac{3}{2}} \|z_{N_1}\|_{L^\infty_t H^{\frac{1}{2}}} \|w_{\sim N}\|_{F^{-\frac{1}{2},\frac{1}{2}}} \\
&\lesssim \|w_{\sim N}\|_{F^{-\frac{1}{2},\frac{1}{2}}} \|z\|_{L^\infty_t H^{\frac{1}{2}}},
\end{aligned}
$$

which is acceptable. Gathering the above estimates we obtain (A-3). □

Gathering Lemmas A.1–A.3 and proceeding as in Section 4B we obtain the local well-posedness in $H^{1/2}$ of (1-3) for $\alpha = 1$. Note that the uniqueness holds in the space $\widetilde{M}^{1/2}_T$.

Acknowledgements

The authors were partially supported by the ANR project GEODISP. They would like to thank the anonymous referee for his careful reading and valuable comments that allowed us to significantly improve the first version of this work.

References

[Abdelouhab et al. 1989] L. Abdelouhab, J. L. Bona, M. Felland, and J.-C. Saut, "Nonlocal models for nonlinear, dispersive waves", *Phys. D* **40**:3 (1989), 360–392. MR 91d:58033 Zbl 0737.35102

[Adams 1975] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics **65**, Academic Press, New York, 1975. MR 56 #9247 Zbl 0314.46030

[Babin et al. 2011] A. V. Babin, A. A. Ilyin, and E. S. Titi, "On the regularization mechanism for the periodic Korteweg-de Vries equation", *Comm. Pure Appl. Math.* **64**:5 (2011), 591–648. MR 2012d:35313 Zbl 1284.35365

[Bona and Smith 1975] J. L. Bona and R. Smith, "The initial-value problem for the Korteweg-de Vries equation", *Philos. Trans. Roy. Soc. London Ser. A* **278**:1287 (1975), 555–601. MR 52 #6219 Zbl 0306.35027

[Bourgain 1993a] J. Bourgain, "Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, II: The KdV-equation", *Geom. Funct. Anal.* **3**:3 (1993), 209–262. MR 95d:35160b Zbl 0787.35098

[Bourgain 1993b] J. Bourgain, "On the Cauchy problem for the Kadomtsev–Petviashvili equation", *Geom. Funct. Anal.* **3**:4 (1993), 315–341. MR 94d:35142 Zbl 0787.35086

[Burq and Planchon 2008] N. Burq and F. Planchon, "On well-posedness for the Benjamin–Ono equation", *Math. Ann.* **340**:3 (2008), 497–542. MR 2009a:35205 Zbl 1148.35074

[Colliander et al. 2004] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Multilinear estimates for periodic KdV equations, and applications", *J. Funct. Anal.* **211**:1 (2004), 173–218. MR 2005a:35241 Zbl 1062.35109

[Edwin and Roberts 1986] P. M. Edwin and B. Roberts, "The Benjamin–Ono–Burgers equation: an application in solar physics", *Wave Motion* **8**:2 (1986), 151–158. MR 87d:76154 Zbl 0587.76193

[Guo 2012] Z. Guo, "Local well-posedness for dispersion generalized Benjamin–Ono equations in Sobolev spaces", *J. Differential Equations* **252**:3 (2012), 2053–2084. MR 2860610 Zbl 1298.35099

[Guo and Wang 2009] Z. Guo and B. Wang, "Global well-posedness and inviscid limit for the Korteweg–de Vries–Burgers equation", *J. Differential Equations* **246**:10 (2009), 3864–3901. MR 2010e:35235 Zbl 1170.35084

[Guo et al. 2011] Z. Guo, L. Peng, B. Wang, and Y. Wang, "Uniform well-posedness and inviscid limit for the Benjamin–Ono-Burgers equation", *Adv. Math.* **228**:2 (2011), 647–677. MR 2012h:35305 Zbl 1234.35072

[Herr et al. 2010] S. Herr, A. D. Ionescu, C. E. Kenig, and H. Koch, "A para-differential renormalization technique for nonlinear dispersive equations", *Comm. Partial Differential Equations* **35**:10 (2010), 1827–1875. MR 2012c:35382 Zbl 1214.35089

[Ionescu and Kenig 2007] A. D. Ionescu and C. E. Kenig, "Global well-posedness of the Benjamin–Ono equation in low-regularity spaces", *J. Amer. Math. Soc.* **20**:3 (2007), 753–798. MR 2008f:35350 Zbl 1123.35055

[Ionescu et al. 2008] A. D. Ionescu, C. E. Kenig, and D. Tataru, "Global well-posedness of the KP-I initial-value problem in the energy space", *Invent. Math.* **173**:2 (2008), 265–304. MR 2009h:35368 Zbl 1188.35163

[Kato 1983] T. Kato, "On the Cauchy problem for the (generalized) Korteweg-de Vries equation", pp. 93–128 in *Studies in applied mathematics*, edited by V. Guillemin, Adv. Math. Suppl. Stud. **8**, Academic Press, New York, 1983. MR 86f:35160 Zbl 0549.34001

[Kenig and Koenig 2003] C. E. Kenig and K. D. Koenig, "On the local well-posedness of the Benjamin–Ono and modified Benjamin–Ono equations", *Math. Res. Lett.* **10**:5–6 (2003), 879–895. MR 2004j:35249 Zbl 1044.35072

[Kenig et al. 1991] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness of the initial value problem for the Korteweg-de Vries equation", *J. Amer. Math. Soc.* **4**:2 (1991), 323–347. MR 92c:35106 Zbl 0737.35102

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR 94h:35229 Zbl 0808.35128

[Koch and Tataru 2007] H. Koch and D. Tataru, "A priori bounds for the 1D cubic NLS in negative Sobolev spaces", *Int. Math. Res. Not.* **2007**:16 (2007), Art. ID rnm053. MR 2010d:35307

[Koch and Tzvetkov 2003] H. Koch and N. Tzvetkov, "On the local well-posedness of the Benjamin–Ono equation in $H^s(\mathbb{R})$", *Int. Math. Res. Not.* **2003**:26 (2003), 1449–1464. MR 2004b:35284 Zbl 1039.35106

[Masmoudi and Nakanishi 2005] N. Masmoudi and K. Nakanishi, "From the Klein–Gordon–Zakharov system to the nonlinear Schrödinger equation", *J. Hyperbolic Differ. Equ.* **2**:4 (2005), 975–1008. MR 2007e:35262 Zbl 1089.35070

[Molinet 2013] L. Molinet, "A note on the inviscid limit of the Benjamin–Ono–Burgers equation in the energy space", *Proc. Amer. Math. Soc.* **141**:8 (2013), 2793–2798. MR 3056569 Zbl 1170.35084

[Molinet and Pilod 2012] L. Molinet and D. Pilod, "The Cauchy problem for the Benjamin–Ono equation in L^2 revisited", *Anal. PDE* **5**:2 (2012), 365–395. MR 2970711 Zbl 1273.35096

[Molinet and Ribaud 2002] L. Molinet and F. Ribaud, "On the low regularity of the Korteweg–de Vries–Burgers equation", *Int. Math. Res. Not.* **2002**:37 (2002), 1979–2005. MR 2003e:35272 Zbl 1031.35126

[Molinet et al. 2001] L. Molinet, J. C. Saut, and N. Tzvetkov, "Ill-posedness issues for the Benjamin–Ono and related equations", *SIAM J. Math. Anal.* **33**:4 (2001), 982–988. MR 2002k:35281 Zbl 0999.35085

[Ott and Sudan 1970] E. Ott and R. N. Sudan, "Damping of solitary waves", *Phys. Fluids* **13**:6 (1970), 1432–1434.

[Ponce 1991] G. Ponce, "On the global well-posedness of the Benjamin–Ono equation", *Differential Integral Equations* **4**:3 (1991), 527–542. MR 92e:35137 Zbl 0732.35038

[Smith 1972] R. Smith, "Nonlinear Kelvin and continental-shelf waves", *J. Fluid Mech.* **52** (1972), 379–391. Zbl 0235.76007

[Tao 2004] T. Tao, "Global well-posedness of the Benjamin–Ono equation in $H^1(\mathbf{R})$", *J. Hyperbolic Differ. Equ.* **1**:1 (2004), 27–49. MR 2005f:35273 Zbl 1055.35104

[Zhou 1997] Y. Zhou, "Uniqueness of weak solution of the KdV equation", *Int. Math. Res. Not.* **1997**:6 (1997), 271–283. MR 98b:35171 Zbl 0883.35105

Received 13 Jan 2015. Revised 24 Apr 2015. Accepted 21 May 2015.

LUC MOLINET: luc.molinet@univ-tours.fr
Laboratoire de Mathématiques et Physique Théorique, CNRS (UMR 7350), Université François Rabelais, Fédération Denis Poisson, Faculté des Sciences et Techniques, Parc de Grandmont, 37200 Tours, France

STÉPHANE VENTO: vento@math.univ-paris13.fr
Laboratoire Analyse, Géométrie et Applications, CNRS (UMR 7539), Université Paris 13, Institut Galilée, 99 avenue J. B. Clément, 93430 Villetaneuse, France

ANALYSIS AND PDE
Vol. 8, No. 6, 2015
dx.doi.org/10.2140/apde.2015.8.1497

msp

ALGEBRAIC ERROR ESTIMATES FOR THE STOCHASTIC HOMOGENIZATION OF UNIFORMLY PARABOLIC EQUATIONS

JESSICA LIN AND CHARLES K. SMART

We establish an algebraic error estimate for the stochastic homogenization of fully nonlinear, uniformly parabolic equations in stationary ergodic spatiotemporal media. The approach is similar to that of Armstrong and Smart in the study of quantitative stochastic homogenization of uniformly elliptic equations.

1. Introduction

We study quantitative stochastic homogenization of equations of the form

$$\begin{cases} u_t^\varepsilon + F(D^2u^\varepsilon, x/\varepsilon, t/\varepsilon^2, \omega) = 0 & \text{in } U_T, \\ u^\varepsilon = g & \text{on } \partial_p U_T, \end{cases} \tag{1-1}$$

where F is a random uniformly elliptic operator, determined by an element ω of some probability space, $U_T := U \times (0, T] \subsetneq \mathbb{R}^{d+1}$ is a compact domain, and $\partial_p U_T$ is the parabolic boundary. Lin [2015] showed that, under suitable hypotheses on the environment (namely stationarity and ergodicity of the operator in space and time), $u^\varepsilon(\,\cdot\,,\cdot\,,\omega)$ converges almost surely to a limiting function u which solves

$$\begin{cases} u_t + \overline{F}(D^2u) = 0 & \text{in } U_T, \\ u = g & \text{on } \partial_p U_T, \end{cases} \tag{1-2}$$

for a uniformly elliptic limiting operator $\overline{F}$ which is independent of ω. Furthermore, a rate of convergence was established under additional quantitative ergodic assumptions. If the environment is strongly mixing with a prescribed logarithmic rate, then the convergence occurs in probability with a logarithmic rate, i.e.,

$$\mathbb{P}\Big[\sup_{U_T} |u^\varepsilon(\,\cdot\,,\cdot\,,\omega) - u(\,\cdot\,,\cdot\,)| \geq f(\varepsilon)\Big] \leq f(\varepsilon) \tag{1-3}$$

MSC2010: primary 35K55; secondary 35K10.
Keywords: quantitative stochastic homogenization, error estimates, parabolic regularity theory.

with $f(\varepsilon) \sim |\log \varepsilon|^{-1}$. In this article, we show that, under the assumption of finite range of dependence, the homogenization occurs in probability with an algebraic rate, i.e., $f(\varepsilon) \sim \varepsilon^{\beta}$.

Background and discussion. For nondivergence form equations in the random setting, the pioneering works establishing the qualitative theory of homogenization (the convergence of $u^{\varepsilon} \to u$) include (but are not limited to) the papers of Papanicolaou and Varadhan [1982] and Yurinskiĭ [1982] for linear, nondivergence form, uniformly elliptic equations, and Caffarelli, Souganidis, and Wang [Caffarelli et al. 2005] for fully nonlinear, uniformly elliptic equations. The study of quantitative stochastic homogenization seeks to establish error estimates for this convergence. For linear, uniformly elliptic equations in nondivergence form, the first results were obtained by Yurinskiĭ [1988; 1991]. Assuming that the environment satisfies an algebraic rate of decorrelation, his works present an algebraic rate of convergence for stochastic homogenization in dimensions $d \geq 5$. In dimensions $d = 3, 4$, the same result holds under the additional assumption of small ellipticity contrast, that is, the ratio of ellipticities is close to 1. In dimension $d = 2$, Yurinskiĭ's results yield a logarithmic rate of convergence.

For fully nonlinear equations, the first quantitative stochastic homogenization result appears in [Caffarelli and Souganidis 2010] for elliptic equations, and the parabolic case with spatiotemporal media was considered in [Lin 2015]. Both of these works obtain logarithmic convergence rates from logarithmic mixing conditions. The approach of both papers was to adapt the obstacle problem method of [Caffarelli et al. 2005] to construct approximate correctors, which play the role of correctors in the random setting. The logarithmic rate appears to be the optimal rate attainable with this approach. This left open the question whether an algebraic rate similar to the results of Yurinskiĭ was attainable in the more general setting of fully nonlinear equations, and for problems in lower dimensions.

In the elliptic setting, this was addressed in [Armstrong and Smart 2014b]. They prove algebraic error estimates in all dimensions for the stochastic homogenization of fully nonlinear, uniformly elliptic equations. The main insight of their work was the introduction of a new subadditive quantity that (1) controls the solutions of the equation and (2) can be studied by adapting the regularity theory of Monge–Ampère equations. Their method does not see the presence of correctors and instead controls solutions indirectly via geometric quantities.

The purpose of this article is to adapt the elliptic strategy to the parabolic spatiotemporal setting, which turns out to be subtle. The approach of [Armstrong and Smart 2014b] was to view the convex envelope of a supersolution as an approximate solution of the Monge–Ampère equation

$$\det D^2 w = 1 \tag{1-4}$$

for w convex and to then use ideas from the regularity theory of (1-4) (namely John's lemma) to control the sublevel sets of w. In the parabolic setting, we will show that the *monotone envelope* of a supersolution of (1-1) is an approximate solution of the analogous Monge–Ampère equation

$$-w_t \det D^2 w = 1 \tag{1-5}$$

for w parabolically convex (convex in space and nonincreasing in time). The equation (1-5) was introduced by Krylov [1976], and then it was pointed out by Tso [1985] that this was the most appropriate parabolic analogue of (1-4). Regularity properties of (1-5) have been studied by Gutiérrez and Huang [1998; 2001] and other parabolic Monge–Ampère equations have been studied by Daskalopoulos and Savin [2012]. In spite of this work, the equation (1-5) is still not as well understood as (1-4). In particular, there is no analogue of John's lemma for sublevel sets of parabolically convex functions. This forced us to develop an alternative approach (which can also be used in the elliptic setting), which replaces John's lemma with a compactness argument.

Assumptions, and statement of the main result. We begin by stating the general assumptions on (1-1) and the precise statement of the main result. We work in the stationary ergodic, spatiotemporal setting. We assume there exists an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\Omega := \{F : \mathbb{S}^d \times \mathbb{R}^{d+1} \to \mathbb{R} \text{ satisfies (F1)–(F4)}\},$$

where (F1)–(F4) will be specified below. In particular, we have $F(X, y, s, \omega) = \omega(X, y, s)$. $\mathcal{F}$ is the Borel σ-algebra on Ω, and we assume that Ω is equipped with a set of measurable, measure-preserving transformations $\tau_{(y',s')} : \Omega \to \Omega$ for each $(y', s') \in \mathbb{R}^{d+1}$. We also assume that $\partial_p U_T$ satisfies a uniform exterior cone condition, which allows us to construct global barriers (see [Crandall et al. 1999] for the precise assumption). Our hypotheses can be summarized as follows:

(F1) *Finite range of dependence*: For $A \subseteq \mathbb{R}^{d+1}$, denote

$$\mathcal{B}(A) := \sigma\{F(\,\cdot\,, y, s, \omega) : (y, s) \in A\},$$

the σ-algebra generated by the operators F defined on A. For $(x_1, t_1), (x_2, t_2) \in \mathbb{R}^{d+1}$, let

$$d[(x_1, t_1), (x_2, t_2)] := (|x_1 - x_2|^2 + |t_1 - t_2|)^{1/2}.$$

For $A, B \subseteq \mathbb{R}^{d+1}$, let

$$d[A, B] := \min\big\{d[(x, t), (y, s)] : (x, t) \in A, (y, s) \in B\big\}. \tag{1-6}$$

The finite range of dependence assumption is:

$$\text{For all random variables } \begin{cases} X : \mathcal{B}(A) \to \mathbb{R}, \\ Y : \mathcal{B}(B) \to \mathbb{R}, \end{cases} \text{ with } d[A, B] \geq 1,\ X, Y \text{ are } \mathbb{P}\text{-independent.} \tag{1-7}$$

(F2) *Stationarity*: For every $(M, \omega) \in \mathbb{S}^d \times \Omega$, where $\mathbb{S}^d$ denotes the space of $d \times d$ symmetric matrices with real entries, and for all $(y', s') \in \mathbb{R}^{d+1}$,

$$F(M, y + y', s + s', \omega) = F(M, y, s, \tau_{(y',s')}\omega).$$

In fact, we only use this hypothesis for $(y', s') \in \mathbb{Z}^{d+1}$.

(F3) *Uniform ellipticity*: For a fixed choice of $\lambda, \Lambda \in \mathbb{R}$ with $0 < \lambda \le \Lambda$, we define Pucci's extremal operators,

$$\mathcal{M}^+(M) = \sup_{\lambda I \le A \le \Lambda I} \{-\mathrm{tr}(AM)\} = -\lambda \sum_{e_i > 0} e_i - \Lambda \sum_{e_i < 0} e_i,$$
$$\mathcal{M}^-(M) = \inf_{\lambda I \le A \le \Lambda I} \{-\mathrm{tr}(AM)\} = -\lambda \sum_{e_i < 0} e_i - \Lambda \sum_{e_i > 0} e_i.$$

We assume that $F(\,\cdot\,, y, s, \omega)$ is uniformly elliptic for each $\omega \in \Omega$, i.e., for all $M, N \in \mathbb{S}^d$ and $(y, s, \omega) \in \mathbb{R}^{d+1} \times \Omega$,

$$\mathcal{M}^-(M - N) \le F(M, y, s, \omega) - F(N, y, s, \omega) \le \mathcal{M}^+(M - N).$$

(F4) *Boundedness and regularity of F*: For every $R > 0$, $\omega \in \Omega$, and $M \in \mathbb{S}^d$ with $|M| \le R$,

$$\{F(M, \cdot\,, \cdot\,, \omega)\} \text{ is uniformly bounded and uniformly equicontinuous on } \mathbb{R}^{d+1},$$

and there exists K_0 such that

$$\operatorname*{ess\,sup}_{\omega \in \Omega} \sup_{(y,s) \in \mathbb{R}^{d+1}} |F(0, y, s, \omega)| < K_0.$$

We also require that there exists a modulus of continuity $\rho[\,\cdot\,]$ and a constant $\sigma > \frac{1}{2}$ such that, for all $(M, y, s, \omega) \in \mathbb{S}^d \times \mathbb{R}^{d+1} \times \Omega$,

$$|F(M, y_1, s_1, \omega) - F(M, y_2, s_2, \omega)| \le \rho\big[(1 + |M|)(|y_1 - y_2| + |s_1 - s_2|)^\sigma\big],$$

where $|\cdot|$ denotes the standard Euclidean norm on $\mathbb{R}^d$ and $\mathbb{R}$ respectively. By applying (F4), we have that

$$\operatorname*{ess\,sup}_{\omega \in \Omega} \sup_{(y,s) \in \mathbb{R}^{d+1}} |F(M, y, s, \omega)| \le C + \Lambda |M| \le C(1 + |M|). \tag{1-8}$$

Equipped with these assumptions, we now state the main result:

Theorem 1.1. *Assume (F1)–(F4), and fix a domain U_T and constant M_0. There exists $C = C(\lambda, \Lambda, d, M_0)$ and a random variable $\mathscr{X} : \Omega \to \mathbb{R}$ with $\mathbb{E}[\exp(\mathscr{X}(\omega))] \le C$ such that, if u^ε solves (1-1), u solves (1-2), and*

$$1 + K_0 + \|g\|_{C^{0,1}(\partial_p U_T)} \le M_0,$$

then, for any $p < d + 2$, there exists a $\beta = \beta(\lambda, \Lambda, d, p) > 0$ such that

$$\sup_{U_T} |u(x, t) - u^\varepsilon(x, t, \omega)| \le C[1 + \varepsilon^p \mathscr{X}(\omega)]\varepsilon^\beta. \tag{1-9}$$

The above theorem implies

$$\mathbb{P}\Big[\sup_{U_T} |u(x, t) - u^\varepsilon(x, t, \omega)| > C\varepsilon^\beta\Big] \le C \exp(-\varepsilon^{-p}) \tag{1-10}$$

for $\beta > 0$ independent of the boundary data. It has recently been shown in the elliptic setting [Armstrong and Smart 2014a; Armstrong and Mourrat 2015; Gloria et al. 2014; Fischer and Otto 2015] that quantitative estimates similar to (1-9) lead to a higher regularity theory at large scales. Although we do not discuss higher regularity results in this article, we are motivated by the recent progress in the elliptic setting to state our results in this form.

Notation and conventions. We mention some general notation and conventions used throughout the paper. The letters λ, Λ, K_0, T, U_T will be used exclusively to refer to the constants stated in the assumptions. In the proofs, the letters c, C will constantly be used as a generic constant which depends on these universal quantities, which may vary line by line, but is precisely specified when needed. We will always denote $\mathbb{S}^d$ as the set of symmetric $d \times d$ matrices with real entries and $\mathbb{M}^d$ as the set of $d \times d$ matrices with real entries. We use the notation $|\cdot|$ to denote a norm on a finite-dimensional Euclidean space ($\mathbb{R}$, $\mathbb{R}^d$, $\mathbb{R}^{d+1}$ or $\mathbb{S}^d$) or the Lebesgue measure on $\mathbb{R}^{d+1}$ and we reserve $\|\cdot\|$ to denote a norm on an infinite-dimensional function space.

We choose to employ the parabolic metric

$$d[(x_1, t_1), (x_2, t_2)] = (|x_1 - x_2|^2 + |t_1 - t_2|)^{1/2}.$$

We point out that this equivalent to the metric

$$d_\infty[(x_1, t_1), (x_2, t_2)] = \max\{|x_1 - x_2|, |t_1 - t_2|^{1/2}\}.$$

We say that $f \in C^{0,\alpha}$ if, for any $(x, t), (y, s) \in \mathbb{R}^{d+1}$,

$$|f(x, t) - f(y, s)| \le \|f\|_{C^{0,\alpha}} d[(x, t), (y, s)]^\alpha.$$

For sets, we use the notation $Q \subseteq \mathbb{R}^{d+1}$ to represent an arbitrary space-time domain, i.e., $Q = Q' \times (t_1, t_2]$, where $Q' \subseteq \mathbb{R}^d$. We define the parabolic boundary by

$$\partial_p Q := (Q' \times \{t = t_1\}) \cup (\partial Q' \times [t_1, t_2)).$$

We use the convention that $\bar{Q} = Q \cup \partial_p Q$, and

$$Q(t) := \{x \in \mathbb{R}^d : (x, t) \in Q\}.$$

We use the conventions

$$\begin{aligned} B_r(\bar{x}, \bar{t}) &= B_r(\bar{x}) \times \{t = \bar{t}\}, \\ \mathcal{B}_r(\bar{x}, \bar{t}) &= \left\{(x, t) \in \mathbb{R}^{d+1} : d[(\bar{x}, \bar{t}), (x, t))] < r\right\}, \\ Q_r(\bar{x}, \bar{t}) &= B_r(\bar{x}) \times (\bar{t} - r^2, \bar{t}]. \end{aligned}$$

In general, B_r, $\mathcal{B}(r)$, and Q_r are used to denote $B_r(0, 0)$, $\mathcal{B}_r(0, 0)$, and $Q_r(0, 0)$, respectively. We point out that $\mathcal{B}_r$ and Q_r are nothing more than the open balls generated by $d[\cdot, \cdot]$ and $d_\infty[\cdot, \cdot]$, respectively.

In addition to these sets, we work with a grid of parabolic cubes which partitions $\mathbb{R}^{d+1}$. The grid boxes take the form

$$G_n = \left[-\tfrac{1}{2}3^n, \tfrac{1}{2}3^n\right)^d \times (0, 3^{2n}].$$

For every $(x,t) \in \mathbb{R}^{d+1}$, we identify the cube

$$G_n(x,t) = \left(3^n \lfloor 3^{-n}x + \tfrac{1}{2} \rfloor, 3^{2n} \lfloor 3^{-2n} t \rfloor\right) + G_n.$$

Outline of the method and the paper. In Section 2, we define the appropriate parabolic analogue of the quantity introduced in [Armstrong and Smart 2014b]. We prove the basic properties of this quantity and describe how it controls solutions from one side. In Section 3, we show how the quantity controls the behavior of solutions from the other side, utilizing the connection with the parabolic Monge–Ampère equation. Here our primary innovation beyond [Armstrong and Smart 2014b] appears.

In Section 4, we construct the effective operator $\overline{F}$ using the asymptotic properties of our quantity and we also construct approximate correctors of (1-1). In Section 5, we obtain a rate of decay on the second moments of this quantity, following closely the analysis of [Armstrong and Smart 2014b]. Finally, in Section 6, we show how the rate on the second moments yields a rate of decay on $|u^\varepsilon - u|$ in probability.

2. A subadditive quantity suitable for parabolic equations

Defining $\boldsymbol{\mu(Q, \omega, \ell, M)}$. We now define the quantity which will be used extensively throughout the rest of the paper. This quantity is a functional which measures the amount a function u bends in space and time. We first recall some geometric objects relevant to the study of parabolic equations and we refer the reader to [Krylov 1976; Wang 1992; Imbert and Silvestre 2012; Gutiérrez and Huang 2001] for general references. We consider a subset $Q \subseteq \mathbb{R}^{d+1}$, a fixed environment $\omega \in \Omega$, $\ell \in \mathbb{R}$, and $M \in \mathbb{S}^d$. We then consider the set

$$S(Q, \omega, \ell, M) = \{u \in C(Q) : u_t + F(M + D^2 u, x, t, \omega) \geq \ell \text{ in } Q\},$$

where the inequality is satisfied in the viscosity sense [Crandall et al. 1992], and, similarly,

$$S^*(Q, \omega, \ell, M) = \{u \in C(Q) : u_t + F(M + D^2 u, x, t, \omega) \leq \ell \text{ in } Q\}.$$

To simplify the notation, we omit parameters when they are assumed to be 0, e.g., $S(Q,\omega)$ refers to the choice $\ell = 0$ and $M = 0$. We say a function u is parabolically convex if $u(\cdot, t)$ is convex for all t and u is nonincreasing in t. For any function u, we define the monotone envelope to be the supremum of all parabolically convex functions lying below u. In particular, Γ^u has the following standard representation formula, which can be taken as the definition:

$$\Gamma^u(x,t) := \sup\{p \cdot x + h : p \cdot y + h \leq u(y,s) \text{ for all } (y,s) \in Q \text{ with } s \leq t\}.$$

We point out that Γ^u depends on the domain Q, however we typically suppress this dependence.

At any point (x_0, t_0), we compute the parabolic subdifferential,

$$\mathcal{P}((x_0,t_0); u) := \left\{(p,h) \subseteq \mathbb{R}^{d+1} : \min_{x \in U, t \leq t_0} u(x,t) - p \cdot x = u(x_0,t_0) - p \cdot x_0 = h\right\},$$

which may be empty.

We then say that, for a domain $Q' \subseteq Q \subseteq \mathbb{R}^{d+1}$,

$$\begin{aligned}\mathcal{P}(Q'; u) &:= \bigcup_{(x_0,t_0)\in Q'} \mathcal{P}((x_0, t_0); u) \\ &= \Big\{(p, h) : \min_{(x,s)\in Q,\, s\le t_0} u(x, s) - p\cdot x = u(x_0, t_0) - p \cdot x_0 = h \text{ for some } (x_0, t_0) \in Q'\Big\}.\end{aligned}$$

We now define the quantity

$$\mu(Q, \omega, \ell, M) := \frac{1}{|Q|} \sup\{|\mathcal{P}(Q; \Gamma^u)| : u \in S(Q, \omega, \ell, M)\}, \tag{2-1}$$

where $|\cdot|$ denotes Lebesgue measure on $\mathbb{R}^{d+1}$.

At this time, we point out some properties of $\mu(Q, \omega)$, which are critical for the analysis which follows:

(1) If u is constant in time, then $Q(t)$ is constant in time. The projection of $\mathcal{P}((x_0, t); u)$ into $\mathbb{R}^d$ is precisely the elliptic subdifferential of the convex envelope of u. We denote the elliptic subdifferential by $\partial\Gamma^u[t](\,\cdot\,;\,\cdot\,)$. This shows that, after an appropriate projection and renormalization, μ as defined in (2-1) reduces to the quantity defined in [Armstrong and Smart 2014b].

(2) This quantity respects the scaling on domains with parabolic scaling. For each $u \in S(G_n, \omega)$, let

$$u_n(x, t) := 3^{-2n} u(3^n x, 3^{2n} t) \in S(G_0, \omega).$$

Under this scaling, if $(p, h) \in \mathcal{P}(G_n; u)$, then $(3^{-n} p, 3^{-2n} h) \in \mathcal{P}(G_0; u_n)$. Thus, we have that

$$|\mathcal{P}(G_n; u)| = 3^{n(d+2)} |\mathcal{P}(G_0; u_n)|.$$

This shows us that, in order to prove statements for $\mu(G_n, \omega)$, it is enough to prove statements for $\mu(G_0, \omega)$ and rescale.

(3) If $w \in C^2(Q)$ is parabolically convex, then $\mathcal{P}((x_0, t_0); w)$ reduces to

$$\mathcal{P}((x, t); w) = (Dw(x, t), w(x, t) - Dw(x, t)\cdot x).$$

If we interpret $\mathcal{P}((\,\cdot\,,\,\cdot\,); w)$ as $\mathcal{P}[w](\,\cdot\,,\,\cdot\,) : \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$, then by a standard computation,

$$\det \mathcal{D}\mathcal{P}[w] = -w_t \det D^2 w,$$

where $\mathcal{D}\mathcal{P}[w] = D_{t,x}\mathcal{P}[w]$. We point out that the right-hand side is precisely the Monge–Ampère operator first introduced in [Krylov 1976; Tso 1985]. Therefore, by applying the area formula [Evans and Gariepy 1992],

$$\frac{1}{|Q|}|\mathcal{P}(Q; w)| = \frac{1}{|Q|}\int_Q \det \mathcal{D}\mathcal{P}[w]\, dx\, dt = \frac{1}{|Q|}\int_Q -w_t \det D^2 w\, dx\, dt.$$

This shows the formal connection between the quantity $|\mathcal{P}(Q; \Gamma^u)|/|Q|$ and the parabolic Monge–Ampère equation. We will explore this connection further in Section 3.

As introduced in [Armstrong and Smart 2014b], we now define $\mu^*(G_n, \omega)$, which will serve as the analogous quantity corresponding to subsolutions. We define the involution operator $\pi(\omega) = \omega^*$ by

$$F(M, x, t, \omega^*) := -F(-M, x, t, \omega) \quad \text{for } (M, x, t, \omega) \in \mathbb{S}^d \times \mathbb{R}^{d+1} \times \Omega.$$

(Recall we assumed Ω is the space of operators F.) We point out that $\pi : \Omega \to \Omega$ is a bijection and $\omega^{**} = \omega$. Moreover, for $u \in C(\overline{Q})$,

$$u_t + F(-M + D^2u, x, t, \omega^*) \geq -\ell \quad \Longleftrightarrow \quad v := -u \text{ solves } v_t + F(M + D^2v, x, t, \omega) \leq \ell$$

in the viscosity sense. Therefore, we define

$$\begin{aligned} \mu^*(Q, \omega, \ell, M) &:= \frac{1}{|Q|} \sup\{|\mathcal{P}(Q; \Gamma^u)| : u \in S(Q, \omega^*, -\ell, -M)\} \qquad (2\text{-}2) \\ &= \mu(Q, \omega^*, -\ell, -M) \\ &= \frac{1}{|Q|} \sup\{|\mathcal{P}(Q; \Gamma^{-u})| : u \in S^*(Q, \omega, \ell, M)\}. \end{aligned}$$

Since $\pi(\omega) = \omega^*$ is an $\mathcal{F}$-measurable function on Ω, we define the pushforward

$$\pi_\# \mathbb{P}(E) := \mathbb{P}[\pi^{-1}(E)].$$

This justifies that $\mu^*(Q, \omega)$ enjoys the analogous properties of $\mu(Q, \omega)$ for subsolutions. Throughout the paper, we will focus on showing results for $\mu(Q, \omega)$; the analogous statements hold for $\mu^*(Q, \omega)$.

Regularity properties of $\mu(Q, \omega)$. First, we show that $\mu(Q, \omega)$ controls the behavior of supersolutions on the parabolic boundary from one side.

Lemma 2.1. *There exists a constant $c_1 = c_1(d)$ such that, for every $\omega \in \Omega$, $(x, t) \in \mathbb{R}^{d+1}$, $n \in \mathbb{Z}$, and $u \in S(G_n(x, t), \omega)$,*

$$\inf_{\partial_p G_n(x,t)} u \leq \inf_{G_n(x,t)} u + c_1 3^{2n} \mu(G_n(x, t), \omega)^{1/(d+1)}. \qquad (2\text{-}3)$$

Proof. Without loss of generality, in light of the scaling of $\mu(\,\cdot\,, \omega)$, it is enough to prove the statement for G_0. Moreover, we assume that $a := \inf_{\partial_p G_0} u - \inf_{G_0} u > 0$. Let $(x_0, t_0) \in G_0$ be such that $u(x_0, t_0) = \inf_{G_0} u$. This implies that, for all $|p| \leq a/\sqrt{d}$ and all $(y, s) \in \partial_p G_0$,

$$\begin{aligned} u(x_0, t_0) - p \cdot x_0 = \inf_{\partial_p G_0} u - a - p \cdot x_0 &\leq u(y, s) - p \cdot y + p \cdot (y - x_0) - a \\ &\leq u(y, s) - p \cdot y + a - a = u(y, s) - p \cdot y, \end{aligned}$$

since $|y - x_0| \leq \sqrt{d}$. This implies that the minimum of the map $(x, t) \to u(x, t) - p \cdot x$ occurs in the interior of G_0. Thus, for all $|p| \leq a/\sqrt{d}$, there exists a choice of h such that $(p, h) \in \mathcal{P}(G_0; u)$.

For each fixed p with $|p| \leq a/\sqrt{d}$, we examine which values of h are included in $\mathcal{P}(G_0; u)$. Recall that

$$h = h(t_0) = \min_{(x,t) \in G_0, t \leq t_0} u(x, t) - p \cdot x.$$

In particular, for each fixed p, the map $h(\cdot):\mathbb{R}\to\mathbb{R}$ is continuous. This implies that $(p,h)\in\mathcal{P}(G_0;u)$ for all $h\in\big[u(x_0,t_0)-p\cdot x_0, \inf_{\partial_p G_0}(u(x,t)-p\cdot x)\big]$.

Combining these observations, this yields that

$$\left\{(p,h):|p|\le\frac{1}{\sqrt{d}}a,\ \inf_{G_0}u-p\cdot x_0\le h\le\inf_{\partial_p G_0}u-p\cdot x\right\}\subseteq\mathcal{P}(G_0;u). \tag{2-4}$$

The left side of (2-4) contains a hypercone in $\mathbb{R}^{d+1}$ with base radius $a/\sqrt{d}$ and height a.

Therefore, we have that, for $c=c(d)$,

$$ca^{d+1}\le|\mathcal{P}(G_0;u)|.$$

Since $\mathcal{P}(G_0;u)\subseteq\mathcal{P}(G_0;\Gamma^u)$, this yields

$$a\le\left(\frac{1}{c}\right)^{\frac{1}{d+1}}\left(\frac{|\mathcal{P}(G_0;\Gamma^u)|}{|G_0|}\right)^{\frac{1}{d+1}}\le c_1\mu(G_0,\omega)^{1/(d+1)}$$

with $c_1=c_1(d)$. □

We now recall several results regarding the regularity of Γ^u. These results and their proofs can be found in [Krylov 1976; Tso 1985; Wang 1992; Imbert and Silvestre 2012].

It is sometimes useful to use an alternative representation formula for the monotone envelope, in terms of its contact points:

Lemma 2.2 [Imbert and Silvestre 2012, Lemma 4.5]. *Γ^u satisfies the alternative representation formula*

$$\Gamma^u(x,t)=\inf\left\{\sum_{i=1}^{d+1}\lambda_i u(x_i,t_i):\sum_{i=1}^{d+1}\lambda_i x_i=x,\ t_i\in[0,t],\ \sum_{i=1}^{d+1}\lambda_i=1,\ \lambda_i\in[0,1]\right\}.$$

In particular, if

$$\Gamma^u(x^0,t^0)=\sum_{i=1}^{d+1}\lambda_i u(x_i^0,t_i^0)\quad\textit{with}\quad\lambda_i>0,$$

then:

- $\Gamma^u(x_i^0,t_i^0)=u(x_i^0,t_i^0)$ *for* $i=1,\dots,d+1$.
- Γ^u *is constant with respect to t and linear with respect to x in the convex set* $\mathrm{co}\{(x_i^0,t^0),(x_i^0,t_i^0)\}_{i=1}^{d+1}$, *the convex hull of* $\{(x_i^0,t^0),(x_i^0,t_i^0)\}_{i=1}^{d+1}$.

As a consequence of this representation formula, it is natural to expect that Γ^u inherits regularity properties of the function u.

Lemma 2.3 [Imbert and Silvestre 2012, Lemma 4.11]. *Suppose that $u_t+\mathcal{M}^+(D^2u)\ge-1$. The function Γ^u is $C^{1,1}$ with respect to x and Lipschitz continuous with respect to t. In particular, $\mathcal{P}[\Gamma^u]:\mathbb{R}^{d+1}\to\mathbb{R}^{d+1}$ is Lipschitz continuous with respect to (x,t).*

In addition, if u is a supersolution to Pucci's equation, it turns out that Γ^u is actually a supersolution to a linear equation almost everywhere:

Lemma 2.4 [Imbert and Silvestre 2012, Lemma 4.12]. *Suppose that $u_t + \mathcal{M}^+(D^2u) \geq -1$. The partial derivatives $(\Gamma^u_t, D^2\Gamma^u)$ satisfy, almost everywhere,*

$$\Gamma^u_t - \lambda\Delta\Gamma^u \geq -1 \quad \textit{in } Q \cap \{u = \Gamma^u\}.$$

We next establish a lemma which shows that, in fact, $|\mathcal{P}(Q; u)| = |\mathcal{P}(Q; \Gamma^u)|$. As previously mentioned, it is immediate that $\mathcal{P}(Q; u) \subseteq \mathcal{P}(Q; \Gamma^u)$ and, thus, $|\mathcal{P}(Q; u)| \leq |\mathcal{P}(Q; \Gamma^u)|$. In order to conclude, it is enough to show the following lemma, which is the parabolic analogue of Lemma 2.4 of [Armstrong and Smart 2014b].

Lemma 2.5. *Let $Q \subseteq \mathbb{R}^{d+1}$ denote an open subset, with $u \in C(Q)$, $(x_0, t_0) \in Q$, and $r > 0$ such that*

$$Q_r(x_0, t_0) \subseteq \{(x, t) \in Q : \Gamma^u(x, t) < u(x, t)\} = \{\Gamma^u < u\}.$$

Then $|\mathcal{P}(Q_r(x_0, t_0); \Gamma^u)| = 0$.

Proof. Without loss of generality, we may assume that $r < 1$. Moreover, by a covering argument, it is enough to show that $|\mathcal{P}(Q_r(x_0, t_0); \Gamma^u)| = 0$ assuming that $Q_{3r}(x_0, t_0) \subseteq \{\Gamma^u < u\}$.

Suppose for the purposes of contradiction that $|\mathcal{P}(Q_r(x_0, t_0); \Gamma^u)| > 0$. Since the measure is positive, by the Lebesgue density theorem almost every $(p, h) \in \mathcal{P}(Q_r(x_0, t_0); \Gamma^u)$ is a density point. We mention that the density theorem still holds for parabolic cylinders and we refer the reader to the appendix of [Imbert and Silvestre 2012] for a proof. We next have the following claim:

Claim. *There exists $(x', t') \in Q_r(x_0, t_0)$ and $(p, h) \in \mathcal{P}((x', t'); \Gamma^u)$ such that (p, h) is a Lebesgue density point of $\mathcal{P}(Q_r(x_0, t_0); \Gamma^u)$ and, also, $p \in \partial\Gamma^u[t'](x')$ is a Lebesgue density point of $\partial\Gamma^u[t'](B_r(x_0))$.*

This follows from applying the Lebesgue density theorem to both $\mathcal{P}(Q_r(x_0, t_0); \Gamma^u)$ and $\partial\Gamma^u[t'](B_r(x_0))$ for some t' with $|\partial\Gamma^u[t'](B_r(x_0))| > 0$. By adding an affine function in space and translating, we may assume that $x_0 = 0$, $t_0 = 0$, $\Gamma^u(x', t') = 0$, and $(p', h') = (0, 0)$.

Since 0 is a Lebesgue density point of $\partial\Gamma^u[t'](B_r)$, for any $\bar{x} \in \partial B_r$ for r sufficiently small there exists a $\bar{p} \in \partial\Gamma^u[t'](B_r) \setminus 0$ such that

$$\bar{p} \cdot \bar{x} \geq \tfrac{3}{4}|\bar{p}|\,|\bar{x}|.$$

Suppose that $\bar{p} \in \partial\Gamma^u[t'](y)$. Since $\Gamma^u(\,\cdot\,, t') \geq 0$ in B_r, this implies that, for any $\alpha \geq 2$,

$$\Gamma^u(\alpha\bar{x}, t') \geq \Gamma^u(y, t') + \bar{p} \cdot (\alpha\bar{x} - y) \geq \alpha\bar{p} \cdot \bar{x} - \bar{p} \cdot y \geq \tfrac{3}{4}\alpha r|\bar{p}| - r|\bar{p}| > 0.$$

This and the monotonicity of Γ^u allows us to conclude that

$$\Gamma^u > 0 \quad \text{on } \{|x| \geq 2r,\ t \leq t'\}.$$

Moreover, we point out that, since $(0, 0)$ is a Lebesgue point of $\mathcal{P}(Q_r; \Gamma^u)$, for each $|x| \leq r < 1$ there exists $(p_2, h_2) \in \mathcal{P}(Q_r; \Gamma^u) \setminus (0, 0)$ such that

$$p_2 \cdot x + h_2 r^2 > \tfrac{3}{4}|(p_2, h_2)|\,|(x, r^2)| > 0.$$

Let $(p_2, h_2) \in \mathcal{P}((y, s); \Gamma^u)$ for $(y, s) \in Q_r$. This implies that, for all $t \le s$ and all $|x| \le r$, since $h_2 \ge 0$ and $r < 1$,

$$\Gamma^u(x, t) \ge p_2 \cdot x + h_2 = p_2 \cdot x + h_2 r^2 + h_2(1 - r^2) > 0.$$

Therefore, for all $t \le -r^2$, we conclude again that $\Gamma^u > 0$. This implies that

$$\Gamma^u > 0 \quad \text{in } (Q \setminus Q_{2r}) \cap \{t \le t'\}.$$

However, since $u > \Gamma^u$ on Q_{3r}, this implies that $u > 0$ on all of $Q \cap \{t \le t'\}$. This contradicts that $\Gamma^u(x', t') = 0$, and hence we have the claim. □

This regularity allows us to establish:

Lemma 2.6. *Assume that $Q \subseteq \mathbb{R}^{d+1}$ is bounded and open, and $u \in C(Q)$ satisfies*

$$u_t + \mathcal{M}^+(D^2 u) \ge -1;$$

then there exists $c_2 = c_2(\lambda, d)$ such that

$$|\mathcal{P}(Q; \Gamma^u)| \le c_2 |\{u = \Gamma^u\} \cap Q|. \tag{2-5}$$

Proof. Given the regularity of Γ^u established by Lemma 2.3, we apply the area formula for Lipschitz functions to conclude that

$$|\mathcal{P}(Q; \Gamma^u)| = \int_Q \det \mathcal{D}\mathcal{P}(\Gamma^u) = \int_{Q \cap \{u = \Gamma^u\}} -\Gamma^u_t \det D^2 \Gamma^u = \lambda^{-d} \int_{Q \cap \{u = \Gamma^u\}} -\Gamma^u_t \det D^2 \lambda \Gamma^u.$$

By applying the geometric–arithmetic mean inequality and Lemma 2.4, we have that

$$\begin{aligned} \lambda^{-d} \int_{Q \cap \{u = \Gamma^u\}} -\Gamma^u_t \det D^2 \lambda \Gamma^u \, dx \, dt &\le c(\lambda, d) \int_{Q \cap \{u = \Gamma^u\}} [-\Gamma^u_t + \lambda \Delta \Gamma^u]^{d+1} \, dx \, dt \\ &\le c \int_{Q \cap \{u = \Gamma^u\}} 1 \, dx \, dt = c |\{u = \Gamma^u\} \cap Q|, \end{aligned}$$

which yields (2-5). □

We next claim that $\lim_{n \to \infty} \mu(G_n, \omega)$ exists almost surely. This will follow by an application of the subadditive ergodic theorem of [Akcoglu and Krengel 1980] to the quantity

$$\sup_{u \in S(G_n, \omega)} |\mathcal{P}(G_n; \Gamma^u)|.$$

We point out that the result of [Akcoglu and Krengel 1980] also holds for cubes with parabolic scaling. In order to verify the hypotheses, we first show a decomposition property of $\mu(\,\cdot\,, \omega)$:

Lemma 2.7. *For each $\omega \in \Omega$, $n \in \mathbb{Z}$, and $m \in \mathbb{N}$,*

$$\mu(G_{n+m}, \omega) \le \fint_{G_{n+m}} \mu(G_n(x, t), \omega) \, dx \, dt. \tag{2-6}$$

Proof. Let $u \in S(G_{n+m}, \omega)$. By applying Lemma 2.6, we have that, for each $(x,t) \in G_{n+m}$,

$$|\mathcal{P}(G_{n+m} \cap \partial_p G_n(x,t); \Gamma^u)| = 0.$$

Therefore,

$$|\mathcal{P}(G_{n+m}; \Gamma^u)| \le \sum_{\{G=G_n(x,t)\subseteq G_{n+m}\}} |\mathcal{P}(G; \Gamma^u)| = \int_{G_{n+m}} \frac{|\mathcal{P}(G_n(x,t); \Gamma^u)|}{|G_n|}\, dx\, dt$$
$$\le \int_{G_{n+m}} \frac{|\mathcal{P}(G_n(x,t); \Gamma^{\tilde{u}})|}{|G_n|}\, dx\, dt,$$

where $\tilde{u} = u\restriction_{G_n(x,t)}$ for $(x,t) \in G_{n+m}$. By taking the supremum of both sides, we have (2-6). □

Lemma 2.7 shows that $\mathbb{E}[\mu(G_n, \omega)]$ is nonincreasing in n. We next show universal bounds for μ.

Lemma 2.8. *There exists $c_3 = c_3(\lambda, \Lambda, d) > 0$ and $c_4 = c_4(\lambda, \Lambda, d) > 0$ such that, for every $\omega \in \Omega$, $n \in \mathbb{Z}$, $M \in \mathbb{S}^d$, and $\ell \in \mathbb{R}$,*

$$c_3 \inf_{(x,t)\in G_n} (F(M,x,t,\omega) - \ell)_+^{d+1} \le \mu(G_n, \omega, \ell, M) \le c_4 \sup_{(x,t)\in G_n} (F(M,x,t,\omega) - \ell)_+^{d+1}. \tag{2-7}$$

Proof. We fix $M \in \mathbb{S}^d$ and, without loss of generality, we assume that $\ell = 0$. By Lemma 2.6, the right inequality holds by scaling and rearranging. To prove the left inequality, we note that, letting

$$\eta := \inf_{(x,t)\in G_n} (F(M,x,t,\omega))_+ \quad \text{and} \quad \varphi(x,t) := -\frac{\eta}{4} t + \frac{\eta}{4d\Lambda}|x|^2,$$

for each $(x,t) \in G_n$ we have

$$\varphi_t + F(M + D^2\varphi, x, t, \omega) \ge \varphi_t + \mathcal{M}^-(D^2\varphi) + F(M,x,t,\omega) = -\frac{\eta}{4} - \frac{\eta}{2} + F(M,x,t,\omega) \ge 0.$$

Therefore, $\varphi \in S(G_n, \omega, M)$, and hence

$$\mu(G_n, \omega, M) \ge \frac{|\mathcal{P}(G_n; \varphi)|}{|G_n|} = \frac{1}{|G_n|} \int -\varphi_t \det D^2\varphi = c_3 \eta^{d+1}.$$ □

In particular, we mention that (2-8) implies

$$c_3 \inf_{(x,t)\in G_n} (F(M,x,t,\omega) - \ell)_+^{d+1} \le \mu(G_n, \omega, \ell, M) \le c_4 [K_0(1+|M|) - \ell]_+^{d+1}. \tag{2-8}$$

Using the previous two lemmas, we establish:

Corollary 2.9. *The limit $\lim_{n\to\infty} \mu(G_n, \omega)$ exists almost surely.*

Proof. We apply the subadditive ergodic theorem to the quantity

$$R(G_n, \omega) := \sup_{u \in S(G_n,\omega)} |\mathcal{P}(G_n; \Gamma^u)|.$$

We note, by the stationarity of $F(\cdot,\cdot,\cdot,\omega)$, it follows that $R(\cdot,\omega)$ is stationary. By Lemma 2.7, Lemma 2.8, and (F4), $R(\cdot,\omega)$ is subadditive on parabolic cubes and bounded almost surely. An application of the subadditive ergodic theorem yields the claim. □

In light of (F1), the limit is a constant almost surely. If $\lim_{n\to\infty}\mu(G_n(x,t),\omega)=0$, then, by (2-3), we obtain a type of comparison principle in the limit. In the next section, we will show that, if the limit is strictly positive, then we obtain control of the growth of an optimizing supersolution.

3. Strict convexity of quasimaximizers

The results in this section are completely deterministic and we suppress all dependencies on the random parameter ω. We show that $|\mathcal{P}(Q;\Gamma^u)|$ yields geometric information about the function $u\in S(Q)$. More specifically, for some $n\le 0$, if $|\mathcal{P}(G_n(x,t);\Gamma^u)|/|G_n|\approx 1$ for all $(x,t)\in G_0$, then the optimizing supersolution for $\mu(G_0)$ is strictly convex. In particular, up to an affine transformation, the optimizing supersolution bends upwards on $\partial_p G_0$.

Formally, if φ is parabolically convex with classical derivatives, then, for n sufficiently small, by the Lebesgue differentiation theorem,

$$-\varphi_t(x,t)\det D^2\varphi(x,t)\approx \fint_{G_n(x,t)} -\varphi_s \det D^2\varphi\,dy\,ds=\frac{|\mathcal{P}(G_n(x,t);\varphi)|}{|G_n|}.$$

Therefore, if $|\mathcal{P}(G_n(x,t);\varphi)|/|G_n|\approx 1$ for all (x,t), this is related to solving the parabolic Monge–Ampère equation $-\varphi_t\det D^2\varphi=1$. This idea originated in [Armstrong and Smart 2014b], where, given an equivalent measure condition for the elliptic subdifferential of the convex envelope, the authors conclude that the optimizing supersolution is strictly convex.

In this article, we first utilize the regularity properties of $u\in S(G_0)$ to show that the time derivatives and Hessian of $w=\Gamma^u$ are uniformly bounded above almost everywhere. In particular, this bound only depends on the ellipticity constants and dimension. Using the structure of (1-5), we then obtain that the time derivative and Hessian are also strictly positive almost everywhere, which allows us to conclude that the solution must be strictly convex. We mention that this approach can also be applied to the elliptic setting of [Armstrong and Smart 2014b] to produce an alternative argument.

We first show that, by using that $u\in S(G_0)$, the monotone envelope Γ^u satisfies a uniform upper bound on the time derivative and Hessian at its contact points. Recall that, by Lemma 2.3, Γ^u is Lipschitz continuous in time and $C^{1,1}$ in space. Therefore, we may represent $(p,h)\in\mathcal{P}((x_0,t_0);\Gamma^u)$ by $(D\Gamma^u(x_0,t_0),u(x_0,t_0)-D\Gamma^u(x_0,t_0)\cdot x_0)\in\mathcal{P}((x_0,t_0);\Gamma^u)$.

Lemma 3.1. *Let $u\in S(G_0)$ and suppose*

$$\frac{|\mathcal{P}(G_{-2}(x,t);\Gamma^u)|}{|G_{-2}|}\le 2 \quad \textit{for all } (x,t)\in G_0. \tag{3-1}$$

There exists $\gamma=\gamma(\lambda,\Lambda,d)$ such that, for all $(x_0,t_0)\in Q_{1/4}(0,1)\cap\{u=\Gamma^u\}$, we have that, for all $(y,s)\in Q_{1/4}(x_0,t_0)$,

$$\Gamma^u(y,s)\le\Gamma^u(x_0,t_0)+D\Gamma^u(x_0,t_0)\cdot(y-x_0)+\gamma. \tag{3-2}$$

Proof. By the monotonicity of Γ^u, it is enough if we can show that for all $y\in B_{1/4}(x_0)$ where $u(x_0,t_0)=\Gamma^u(x_0,t_0)$,

$$\Gamma^u\big(y,t_0-\tfrac{1}{16}\big)\le\Gamma^u(x_0,t_0)+D\Gamma^u(x_0,t_0)\cdot(y-x_0)+\gamma. \tag{3-3}$$

We proceed by contradiction. Let $w := \Gamma^u$ be defined in G_0. Assume that there exists a point (x_0, t_0) such that

$$\sup_{B_{1/4}(x_0,t_0)} w\bigl(\,\cdot\,, t_0 - \tfrac{1}{16}\bigr) > w(x_0, t_0) + Dw \cdot (y - x_0) + \gamma, \tag{3-4}$$

with γ to be chosen. Without loss of generality, by adding an affine function, we may assume that $(x_0, t_0) = (0, 1)$ and $\Gamma^u(x_0, t_0) = D\Gamma^u(x_0, t_0) = 0$.

Choose $\bar{y} \in \bar{B}_{1/4}$ so that

$$w\bigl(\bar{y}, \tfrac{15}{16}\bigr) := \max_{\bar{B}_{1/4}} w\bigl(\,\cdot\,, \tfrac{15}{16}\bigr).$$

By (3-4),

$$w\bigl(\bar{y}, \tfrac{15}{16}\bigr) > \gamma.$$

Since $w\bigl(\,\cdot\,, \tfrac{15}{16}\bigr)$ is convex, and using the definition of $\bar{y}$, this implies that

$$w\bigl(z, \tfrac{15}{16}\bigr) > \gamma \quad \text{for all } z \text{ such that } z \cdot \bar{y} \geq |\bar{y}|^2.$$

In particular, let $\Theta := \bigl\{\bigl(z, \tfrac{15}{16}\bigr) : z \in B_{1/2},\ z \cdot \bar{y} \geq |\bar{y}|^2\bigr\}$.

Let $\mathcal{Q} := B_{1/2} \times \bigl(\tfrac{15}{16}, 1\bigr]$. We claim there exists a test function $\varphi \in C^2(\mathcal{Q})$ which satisfies

$$\begin{cases} \varphi_t + \mathcal{M}^-(D^2\varphi) \geq 0 & \text{in } \mathcal{Q}, \\ \varphi \geq -\chi_\Theta & \text{on } \partial_p \mathcal{Q}, \end{cases} \tag{3-5}$$

and $\min \varphi(\,\cdot\,, 1) \leq -c$ for some universal constant c. First, by approximating $-\chi_\Theta$ by a smooth function from above and applying the Evans–Krylov theorem [Krylov 1982], there exists a supersolution which is C^2 satisfying the boundary conditions of (3-5).

By the strong maximum principle, there exists a nonconstant solution such that $\min \varphi(\,\cdot\,, 1) \leq -c$. Moreover, by compactness, this c can be chosen universally for all $(x_0, t_0) \in Q_{1/4}(0, 1)$ by a standard covering argument. This implies that $u + \gamma\varphi$ satisfies

$$\begin{cases} (u + \gamma\varphi)_t + F(D^2(u + \gamma\varphi), x, t) \geq 0 & \text{in } \mathcal{Q}, \\ u + \gamma\varphi \geq 0 & \text{on } \partial_p \mathcal{Q}, \\ \min_{\mathcal{Q}} (u + \gamma\varphi)(\,\cdot\,, 1) \leq -c\gamma. \end{cases}$$

By a similar estimate as in Lemma 2.1, this implies that $|\mathcal{P}(\mathcal{Q})| \geq c\gamma^{d+1}$. Therefore, if we consider covering $\mathcal{Q}$ with a collection of $G_{-2}(x, t) \subseteq G_0$, then

$$c\gamma^{d+1} \leq \sum_{G_{-2}(x,t) \subset G_0} |\mathcal{P}(G_{-2}(x, t))| \leq 2|G_0|.$$

Choosing γ sufficiently large, depending only on λ, Λ, and d, we obtain a contradiction. Therefore, (3-2) holds. □

By rescaling Lemma 3.1, we actually have that if, for all $(x, t) \in G_0$,

$$\frac{|\mathcal{P}(G_n(x, t); u)|}{|G_n|} \leq 2 \quad \text{and} \quad 3^n \leq \tfrac{1}{4} r,$$

then, for any point such that $u(x_0, t_0) = \Gamma^u(x_0, t_0)$, for all $(y, s) \in Q_r(x_0, t_0)$,

$$\Gamma^u(y, s) \leq \Gamma^u(x_0, t_0) + D\Gamma^u(x_0, t_0) \cdot (y - x_0) + \gamma r^2. \tag{3-6}$$

By sending $r \to 0$, this implies that $\Gamma^u_t \leq \gamma$ and $D^2\Gamma^u \leq \gamma \,\mathrm{Id}$ at all contact points where $u = \Gamma^u$. By the construction of the monotone envelope (in particular, Lemma 2.2), this implies that $\Gamma^u_t \leq \gamma$ and $D^2\Gamma^u \leq \gamma \,\mathrm{Id}$ everywhere in G_0. The proof is identical to the proof of Lemma 2.3, which can be found in [Imbert and Silvestre 2012]. We choose to omit it since it follows verbatim.

We highlight that, unlike Lemma 2.3, the upper bound on the time derivatives and Hessian of Γ^u will be *independent of* K_0. An observation of [Armstrong and Smart 2014b] is that it does not seem feasible to obtain an algebraic rate if these upper bounds depend on K_0. Recall that our goal is to establish an estimate which controls supersolutions from the other side of Lemma 2.1. Since we plan on performing quantitative analysis, it is important that our estimate is *scale-invariant*. If our estimate depended on K_0 then, by (F4), the estimate would depend upon the scaling. In general, the upper bounds on the time derivative and the Hessian are controlled by the quantity $\mu(G_n(x, t))$. In light of (3-1), this is enough to conclude that γ is independent of K_0.

We next show that these upper bounds are actually enough to conclude strict convexity.

Lemma 3.2. *There exists $c_5 = c_5(\lambda, \Lambda, d) > 0$ such that, for every $\varepsilon > 0$, there exists $n_1 = n_1(\varepsilon, d) < 0$ such that, if $u \in S(G_0)$ and $n \leq n_1$ satisfies*

$$1 \leq \frac{|\mathcal{P}(G_n(x, t); \Gamma^u)|}{|G_n|} \leq 2 \quad \textit{for all } (x, t) \in G_0, \tag{3-7}$$

then, for all $(x_0, t_0) \in Q_{1/4}(0, 1) \cap \{u = \Gamma^u\}$ and all $(y, s) \in Q_{1/4}(x_0, t_0)$,

$$\Gamma^u(y, s) \geq \Gamma^u(x_0, t_0) + D\Gamma^u(x_0, t_0) \cdot (y - x_0) + c_5(t_0 - s + |y - x_0|^2) - \varepsilon. \tag{3-8}$$

Proof. Fix $\varepsilon > 0$. Suppose for the purposes of contradiction that (3-8) does not hold. Therefore, there exists a sequence of $(u_n, \hat{y}_n, \hat{s}_n) \in S(G_0) \times G_0$ such that u_n satisfies (3-7) for n and u_n violates (3-8) at $(\hat{y}_n, \hat{s}_n)$. Using the convention that $w_n := \Gamma^{u_n}$ and, without loss of generality, assuming that $w_n \geq 0$ in G_0 and $w_n(0, 1) = 0$ for each n, this amounts to

$$w_n(\hat{y}_n, \hat{s}_n) < c(\hat{s}_n + |\hat{y}_n|^2) - \varepsilon \tag{3-9}$$

for c to be chosen.

By (3-6) and (3-2), the family $\{w_n\}$ is equicontinuous and uniformly bounded in $Q_{1/4}(0, 1)$. By the Arzelà–Ascoli theorem, this implies that there exists a subsequence converging uniformly to a limiting function w, with w satisfying

$$-w_t \leq \gamma \quad \text{and} \quad D^2 w \leq \gamma \,\mathrm{Id} \quad \text{almost everywhere.}$$

By the Lebesgue differentiation theorem and (3-7), w also satisfies

$$1 \leq -w_t \det D^2 w \leq 2 \quad \text{almost everywhere.}$$

Therefore, this yields that $-w_t \geq 1/\gamma^d$, and $\det D^2 w \geq (1/\gamma)\,\mathrm{Id}$ almost everywhere. Since $D^2 w \leq \gamma\,\mathrm{Id}$, this yields that there exists a constant $c_\gamma = c(\gamma, d)$ such that $D^2 w \geq c_\gamma\,\mathrm{Id}$.

Consider that, by (3-9), since $(\hat{y}_n, \hat{s}_n) \in G_0$, there exists a subsequence converging to a point $(\hat{y}, \hat{s}) \in G_0$ satisfying

$$w(\hat{y}, \hat{s}) < c(\hat{s} + |\hat{y}|^2) - \varepsilon.$$

However, for c chosen appropriately in terms of γ, this contradicts $-w_t \geq 1/\gamma^d$, $D^2 w \geq (1/\gamma)\,\mathrm{Id}$ almost everywhere. □

Finally, we show that this implies that u will also be strictly convex on the parabolic boundary.

Theorem 3.3. *Let $u \in S(G_1)$. There exist constants $c_6 = c_6(\lambda, \Lambda, d)$ and $n_1 = n_1(d) < 0$ such that, if $n \leq n_1$ satisfies*

$$1 \leq \frac{|\mathcal{P}(G_n(x,t);\Gamma^u)|}{|G_n|} \leq \mu(G_n(x,t)) \leq 1 + 3^{n(d+2)} \quad \textit{for all } (x,t) \in G_1, \tag{3-10}$$

then there exists a point $(x_0, t_0) \in \{u = \Gamma^u\} \cap G_n(0, 9)$ and $(p_0, h_0) \in \mathcal{P}((x_0, t_0); \Gamma^u)$ such that

$$u(x,t) \geq p_0 \cdot x + h_0 + c_6 \quad \textit{for all } \{t \leq t_0\} \cap G_1 \setminus G_0(0,9). \tag{3-11}$$

Proof. In order to prove (3-11), it is enough to obtain a lower bound on $\inf_{\partial_p G_0(0,9)} \Gamma^u(\,\cdot\,, t)$ for $t \leq t_0$. We claim there exists $(x_0, t_0) \in G_n(0, 9)$ such that $u(x_0, t_0) = \Gamma^u(x_0, t_0)$. By (3-10), for any $(y, s) \in G_n(0, 9)$,

$$\begin{aligned} 1 &\leq \int_{G_0(0,9)} \frac{|\mathcal{P}(G_n(x,t);\Gamma^u)|}{|G_n|}\,dx\,dt \\ &= |\mathcal{P}(G_n(y,s);\Gamma^u)| + \int_{G_0(0,9)\setminus G_n(y,s)} \frac{|\mathcal{P}(G_n(x,t);\Gamma^u)|}{|G_n|}\,dx\,dt \\ &\leq |\mathcal{P}(G_n(y,s);\Gamma^u)| + (1 - 3^{n(d+2)})(1 + 3^{n(d+2)}). \end{aligned}$$

This shows that $|\mathcal{P}(G_n(y,s);\Gamma^u)| > 0$ for any $(y,s) \in G_0$, which implies, by Lemma 2.6, that

$$|G_n(0,9) \cap \{u = \Gamma^u\}| > 0.$$

Let $(x_0, t_0) \in G_n(0,9) \cap \{u = \Gamma^u\}$ and consider $(p_0, h_0) \in \mathcal{P}((x_0, t_0); \Gamma^u)$. Let $\tilde{u}(x,t) = u(x,t) - p_0 \cdot x - h_0$. Then $\tilde{u} \in S(G_0(0,9))$ and $\tilde{u}(x_0, t_0) = \Gamma_{\tilde{u}}(x_0, t_0) = 0$. Moreover, we have that $(0, 0) \in \mathcal{P}((x_0, t_0); \Gamma^{\tilde{u}})$ and $\Gamma^{\tilde{u}} \geq 0$ for all $(x,t) \in G_0(0,9) \cap \{t \leq t_0\}$.

By Lemma 3.2, letting $\varepsilon = \frac{1}{2}c_5$, since $Q_{1/4}(x_0, t_0) \subset G_0(0,9)$, this implies that, on $\partial_p G_0(0,9)$,

$$u(x,t) \geq \Gamma^u(x,t) \geq \tfrac{1}{2}c_5.$$

Defining $c_6 := \frac{1}{2}c_5$ completes the proof. □

For convenience, we also provide a rescaled version of (3-11) which will be used extensively later in the paper. Let $u \in S(G_{m+n+1})$. Choose $n \leq n_1$ so that

$$\alpha \leq \frac{|\mathcal{P}(G_n(x,t);\Gamma^u)|}{|G_n|} \leq \mu(G_n(x,t)) \leq (1 + 3^{n(d+2)})\alpha \quad \text{for all } (x,t) \in G_{m+n+1}.$$

There exists a point $(x_0, t_0) \in \{u = \Gamma^u\} \cap G_n(0, 3^{2(m+n+1)})$ and $(p_0, h_0) \in \mathcal{P}((x_0, t_0); \Gamma^u)$ such that

$$u(x,t) \geq p_0 \cdot x + h_0 + c_6 \alpha^{1/(d+1)} 3^{2(m+n)} \quad \text{for all } \{t \leq t_0\} \cap G_{m+n+1} \setminus G_{m+n}(0, 3^{2(m+n+1)}). \tag{3-12}$$

4. The construction of $\bar{F}$ and the construction of approximate correctors

We now define the homogenized operator $\bar{F} : \mathbb{S}^d \to \mathbb{R}$. In addition, we show how one can obtain "approximate correctors" as in [Lin 2015] using the quantity μ. For each $M \in \mathbb{S}^d$, we say that w^ε is an approximate corrector of (1-1) if there exists w^ε satisfying

$$\begin{cases} w^\varepsilon_t + F(M + D^2 w^\varepsilon, x, t, \omega) = \bar{F}(M) & \text{in } Q_{1/\varepsilon}, \\ w^\varepsilon = 0 & \text{on } \partial_p Q_{1/\varepsilon}, \end{cases} \tag{4-1}$$

with $\|\varepsilon^2 w^\varepsilon\|_{L^\infty(Q_{1/\varepsilon})} \to 0$ as $\varepsilon \to 0$. Once w^ε exists, the qualitative homogenization (the convergence $u^\varepsilon \to u$ $\mathbb{P}$-a.s.) follows by a standard perturbed test function argument [Evans 1992], as shown in [Lin 2015]. In particular, the uniform ellipticity of $\bar{F}$ follows from the existence of approximate correctors.

Identifying $\bar{F}$. We identify $\bar{F}(M)$ for each fixed $M \in \mathbb{S}^d$. First, we establish a lemma which states that μ is Lipschitz continuous with respect to the right-hand side ℓ.

Lemma 4.1. *There exists $C(\lambda, \Lambda, d, M, K_0) > 0$ such that*

$$0 \geq \mu(Q, \omega, \ell + s, M) - \mu(Q, \omega, \ell, M) \geq -C|Q|s \tag{4-2}$$

for all $s \in [0, 1]$.

Proof. Since $S(Q, \omega, \ell+s, M) \subseteq S(Q, \omega, \ell, M)$, the left inequality follows from the comparison principle for viscosity solutions. To obtain the right inequality, let $u \in S(Q, \omega, \ell, M)$ and define $u^s(x,t) := u(x,t) + st$, which lies in $S(Q, \omega, \ell + s, M)$. Let w^s denote the monotone envelope of u^s. We note that $|w^s_t|, |D^2 w^s| \leq C(K_0, \ell + s, M)$ on the contact set $\{u^s = w^s\}$, by Lemma 2.3 and Lemma 2.6. Therefore, by the area formula, this implies that

$$\begin{aligned} |\mathcal{P}(Q; w^s)| &= \int_{\{u^s = w^s\} \cap Q} -u^s_t \det D^2 u^s \, dx, \\ &\geq \int_{\{u=w\} \cap \{u_t \leq -s\} \cap Q} -u^s_t \det D^2 u^s \, dx, \\ &\geq \int_{\{u=w\} \cap Q} -u_t \det D^2 u - Cs|Q| \\ &= |\mathcal{P}(Q; w)| - Cs|Q|. \end{aligned}$$

By taking the supremum over $u \in S(Q, \omega, \ell, M)$, this yields (4-2). □

Lemma 4.2. *Let $M \in \mathbb{S}^d$. For every $n \in \mathbb{N}$, the map*

$$\ell \to \mathbb{E}[\mu(G_n, \omega, \ell, M)] \quad \textit{is continuous and nonincreasing.}$$

Similarly, the map

$$\ell \to \mathbb{E}[\mu^*(G_n, \omega, \ell, M)] \quad \textit{is continuous and nondecreasing.}$$

In addition, there exists $\bar{\ell}(M) \in \mathbb{R}$ such that, $\mathbb{P}$-a.s. in ω,

$$\lim_{n\to\infty} \mu(G_n, \omega, \bar{\ell}(M), M) = \lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega, \bar{\ell}(M), M)] = \lim_{n\to\infty} \mathbb{E}[\mu^*(G_n, \omega, \bar{\ell}(M), M)] = \lim_{n\to\infty} \mu^*(G_n, \omega, \bar{\ell}(M), M). \tag{4-3}$$

Proof. The Lipschitz continuity and monotonicity follow from Lemma 4.1. By (2-8), $\mathbb{E}[\mu(G_n, \omega, \ell)] = 0$ for all $\ell \geq K_0(1+|M|)$. In particular, this implies that

$$\lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega, \ell)] = 0 \quad \text{for all } \ell \geq K_0(1+|M|).$$

Similarly,

$$\lim_{n\to\infty} \mathbb{E}[\mu^*(G_n, \omega, \ell)] = 0 \quad \text{for all } \ell \leq -K_0(1+|M|).$$

Using the monotonicity in ℓ and (2-8), there exists a choice of $\bar{\ell}$ such that $\lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega, \bar{\ell})] = \lim_{n\to\infty} \mathbb{E}[\mu^*(G_n, \omega, \bar{\ell})]$. The outer equalities of (4-3) hold in light of the ergodicity assumption (F1) and the subadditive ergodic theorem. □

Using Lemma 4.2, we define

$$\bar{F}(M) := \bar{\ell}(M). \tag{4-4}$$

We will now show that $\bar{F}(M)$ agrees with the effective operator constructed in [Lin 2015] and thus the uniqueness follows. To do this, it is enough to show that solutions w^ε of (4-1) exist and satisfy the desired limiting behavior.

A qualitative homogenization argument. The construction of approximate correctors (4-1) follows in two steps. First we show that, for any $M \in \mathbb{S}^d$, it is impossible for $E(\bar{\ell}(M), M) := \lim_{n\to\infty} \mu(G_n, \omega, \bar{\ell}(M), M)$ and $E^*(\bar{\ell}, M) := \lim_{n\to\infty} \mu^*(G_n, \omega, \bar{\ell}(M), M)$ to both be positive. Applying Lemma 2.1 allows us to conclude.

For convenience, we provide a precise statement of the Harnack inequality for parabolic equations, as can be found in [Wang 1992; Imbert and Silvestre 2012]. We will use the notation of this theorem in the future.

Theorem 4.3 (Harnack inequality). *Let u be nonnegative with $-|f| \leq u_t + \mathcal{M}^+(D^2u) \leq |f|$. Then there exists a universal $C = C(\lambda, \Lambda, d)$ such that*

$$\sup_{\tilde{Q}} u \leq C\Big(\inf_{Q_{\rho^2}} u + \|f\|_{L^{d+1}(Q_1)}\Big),$$

where $\tilde{Q} := B_{\rho^2/(2\sqrt{2})} \times \big(-\rho^2 + \frac{3}{8}\rho^4, -\rho^2 + \frac{1}{2}\rho^4\big) \subseteq Q_1$ and $\rho = \rho(\lambda, \Lambda, d)$.

The Harnack inequality implies that E and E^* must vanish when they are equal:

Lemma 4.4. *Fix $M \in \mathbb{S}^d$. If $\ell \in \mathbb{R}$ is such that*

$$\lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega, \ell, M)] = E(\ell, M) = E^*(\ell, M) = \lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega^*, -\ell, M)], \tag{4-5}$$

then $E(\ell, M) = E^*(\ell, M) = 0$.

Proof. We drop the dependence on M since it is fixed throughout the proof. Suppose that both $E(\ell) = E^*(\ell) := \alpha > 0$. By the subadditive ergodic theorem, there exists a choice of m sufficiently large such that, for all $(x, t) \in G_{m+n}$, with n large to be chosen,

$$\tfrac{1}{2}\alpha \leq \frac{|\mathcal{P}(G_m(x,t);\Gamma^u)|}{|G_m|} \leq \mu(G_m, \omega, \ell) \leq 2\alpha.$$

Without loss of generality, we assume that $m = 0$. By Theorem 3.3, rescaled, choosing n sufficiently large, and after an affine transformation, there exists a function u such that

$$u_t + F(D^2u, x, t, \omega) = \ell \quad \text{in } G_n(0, 3^{2(n+1)}) \tag{4-6}$$

and $(x_0, t_0) \in G_0(0, 3^{2(n+1)})$ such that

$$u \geq u(x_0, t_0) + C3^{2n}\alpha^{1/(d+1)} \quad \text{on } \partial_p G_n(0, 3^{2(n+1)}) \cap \{t \leq t_0\} \tag{4-7}$$

and

$$\inf_{G_n(0,3^{2(n+1)})\cap\{t\leq t_0\}} u = \inf_{G_0(0,3^{2(n+1)})\cap\{t\leq t_0\}} u = u(x_0, t_0) = 0.$$

This is done by extracting $u' \in S(G_{n+1}, \omega)$ such that (3-11) holds. Upon an affine transformation and solving (4-6) with $u = u'$ on $\partial_p G_n(0, 3^{2(n+1)})$, we have the claim. Similarly, there exists u^* satisfying

$$u^*_t + F(D^2u^*, x, t, \omega^*) = -\ell \quad \text{in } G_n(0, 3^{2(n+1)}) \tag{4-8}$$

and, for some $(x^*_0, t^*_0) \in G_0(0, 3^{2(n+1)})$,

$$u^* \geq u^*(x_0, t_0) + C3^{2n}\alpha^{1/(d+1)} \quad \text{on } \partial_p G_n(0, 3^{2(n+1)}) \cap \{t \leq t^*_0\} \tag{4-9}$$

and

$$\inf_{G_n(0,3^{2(n+1)})\cap\{t\leq t^*_0\}} u^* = \inf_{G_0(0,3^{2(n+1)})\cap\{t\leq t^*_0\}} u^* = u^*(x^*_0, t^*_0) = 0.$$

Let $\bar{t} = \min\{t_0, t^*_0\}$. Notice that $w := u + u^*$ satisfies

$$w_t + \mathcal{M}^+(D^2w) \geq u_t + u^*_t + F(D^2u, x, t, \omega) + F(D^2u^*, x, t, \omega^*) = 0 \quad \text{in } G_n(0, 3^{2(n+1)})$$

and

$$w \geq C3^{2n}\alpha^{1/(d+1)} \quad \text{on } \partial_p G_n(0, 3^{2(n+1)}) \cap \{t \leq \bar{t}\}.$$

By the Alexandrov–Backelman–Pucci–Krylov–Tso estimate [Wang 1992; Imbert and Silvestre 2012], this implies that

$$w \geq C3^{2n}\alpha^{1/(d+1)} \quad \text{in } G_n(0, 3^{2(n+1)}) \cap \{t \leq \bar{t}\}. \tag{4-10}$$

Let s be defined as the smallest integer such that $\rho^2 3^s \geq \sqrt{d}$, where ρ is defined in the Harnack inequality (Theorem 4.3). We may assume that $s \leq n$, by choosing n larger if necessary. We observe that, in $G_s(0, 3^{2(n+1)})$, u and u^* also each satisfy

$$u_t + \mathcal{M}^+(D^2u) \geq -|\ell| - K_0 \quad \text{and} \quad K_0 + |\ell| \geq u_t + \mathcal{M}^-(D^2u),$$
$$u^*_t + \mathcal{M}^+(D^2u^*) \geq -|\ell| - K_0 \quad \text{and} \quad |\ell| + K_0 \geq u^*_t + \mathcal{M}^-(D^2u^*).$$

Since $\inf_{G_0(0,3^{2(n+1)})} u = \inf_{G_0(0,3^{2(n+1)})} u^* = 0$, and

$$G_0(0, 3^{2(n+1)}) \subseteq Q_{\rho^2 3^s}(0, 3^{2(n+1)})$$

by our choice of s, this implies, by the Harnack inequality, that there exists $C = C(\lambda, \Lambda, d, \ell, K_0)$ such that

$$\sup_{\tilde{Q}} u \le C3^{2s} \quad \text{and} \quad \sup_{\tilde{Q}} u^* \le C3^{2s},$$

where $\tilde{Q} \subseteq G_s(0, 3^{2(n+1)})$ is a rescaled version of the $\tilde{Q}$ defined in Theorem 4.3. Thus, there exists $C = C(\lambda, \Lambda, d, \ell, K_0) > 0$ such that

$$w \le C3^{2s} \quad \text{in } \tilde{Q} \subseteq G_s(0, 3^{2(n+1)}).$$

By choosing n sufficiently large, depending on ℓ, K_0, and α, we obtain a contradiction with (4-10). Therefore, $\alpha = 0$. □

We next show that w^ε solving (4-1) has the desired decay with this definition of $\bar{F}(M)$. Letting $\varepsilon = 3^{-n}$, we relabel (4-1) as

$$\begin{cases} w_t^n + F(M + D^2 w^n, x, t, \omega) = \bar{F}(M) & \text{in } G_n, \\ w^n = 0 & \text{on } \partial_p G_n, \end{cases} \tag{4-11}$$

and we want to show that $\|3^{-2n} w^n\|_{L^\infty(G_n)} \to 0$ as $n \to \infty$.

Consider that, since $E(\bar{F}(M), M) = E^*(\bar{F}(M), M) = 0$, this implies that, almost surely,

$$\lim_{n\to\infty} \mu(G_n, \omega) = 0 = \lim_{n\to\infty} \mu^*(G_n, \omega).$$

By Lemma 2.1 and (4-11), this implies that

$$\begin{aligned} 0 &\le \inf_{G_n} 3^{-2n} w^n + c_1 \mu(G_n, \omega)^{1/(d+1)} \\ \text{and} \quad 0 &\ge \sup_{G_n} 3^{-2n} w^n - c_1 \mu^*(G_n, \omega)^{1/(d+1)}. \end{aligned}$$

Taking $n \to \infty$, this yields

$$\lim_{n\to\infty} \|3^{-2n} w^n\|_{L^\infty(G_n)} \le \lim_{n\to\infty} \max\{\mu(G_n, \omega)^{1/(d+1)}, \mu^*(G_n, \omega)^{1/(d+1)}\} = 0, \tag{4-12}$$

as desired.

5. A rate of decay on the second moments

In this section, we obtain a rate of decay on the second moments of μ. The approach of this section closely follows that of [Armstrong and Smart 2014b]. As before, we suppress the dependence on M. We simplify the notation by adopting the following conventions. Let

$$E_n(\ell) = \mathbb{E}[\mu(G_n, \omega, \ell)] \quad \text{and} \quad E_n^*(\ell) = \mathbb{E}[\mu^*(G_n, \omega, \ell)] = \mathbb{E}[\mu(G_n, \omega^*, -\ell)].$$

Also, let

$$J_n(\ell) = \mathbb{E}[\mu(G_n, \omega, \ell)^2] \quad \text{and} \quad J_n^*(\ell) = \mathbb{E}[\mu^*(G_n, \omega, \ell)^2] = \mathbb{E}[\mu(G_n, \omega^*, -\ell)^2].$$

Our next lemma shows that, if the variance of μ and μ^* are not decaying, then their expectations must be close to zero. The proof resembles the argument for Lemma 4.4, but avoids the dependence on K_0.

Lemma 5.1. *Suppose that there exists $m, n \in \mathbb{N}$ and $\eta, \gamma > 0$ such that*

$$0 < J_m(\ell - \gamma) \leq (1+\eta) E^2_{m+n}(\ell - \gamma) \tag{5-1}$$

and

$$0 < J^*_m(-\ell + \gamma) \leq (1+\eta) E^{*2}_{m+n}(-\ell + \gamma). \tag{5-2}$$

Then there exists $n_0 = n_0(\lambda, \Lambda, d)$ and $\eta_0 = \eta_0(\lambda, \Lambda, d)$ such that, for all $n \geq n_0$ and all $\eta \leq \eta_0$,

$$J_{m+n}(\ell - \gamma) + J^*_{m+n}(-\ell + \gamma) \leq C\gamma^{2(d+1)}. \tag{5-3}$$

Proof. Without loss of generality, we assume that $\ell = 0$ and $m = 0$. First, we claim that there exists a choice of environment ω such that $\mu(G_n, \omega)$ and $\mu(G_0(x,t), \omega)$ are approximately constant for all $(x,t) \in G_n$.

Fix $\delta > 0$. There exists $\eta = \eta(\delta)$ such that, if (5-1) and (5-2) hold for this η, there exists an ω such that, for all $(x,t) \in G_n$,

$$(1-\delta) E_n(-\gamma) \leq \mu(G_n, \omega, -\gamma) \leq \mu(G_0(x,t), \omega, -\gamma) \leq (1+\delta) E_n(-\gamma) \tag{5-4}$$

and, similarly for the lower quantity,

$$(1-\delta) E^*_n(\gamma) \leq \mu^*(G_n, \omega, \gamma) \leq \mu^*(G_0(x,t), \omega, \gamma) \leq (1+\delta) E^*_n(\gamma). \tag{5-5}$$

Applying Chebyshev's inequality, we have that, for any $(x,t) \in G_n$,

$$\begin{aligned}
\mathbb{P}\big[\mu(G_0(x,t), \omega, -\gamma) \geq (1+\delta) E_n(-\gamma)\big] &\leq \mathbb{P}\big[\mu(G_0(x,t), \omega, -\gamma) - E_n(-\gamma) \geq \delta E_n(-\gamma)\big] \\
&\leq \mathbb{P}\big[[\mu(G_0(x,t), \omega, -\gamma) - E_n(-\gamma)]^2 \geq \delta^2 E^2_n(-\gamma)\big] \\
&\leq \frac{1}{\delta^2 E^2_n(-\gamma)} \mathbb{E}\big[[\mu(G_0(x,t), \omega, -\gamma) - E_n(-\gamma)]^2\big] \\
&\leq \frac{1}{\delta^2 E^2_n(-\gamma)} [J_0(-\gamma) - E^2_n(-\gamma)] \\
&\leq \eta\delta^{-2},
\end{aligned}$$

where the last inequality follows from (5-1).

Similarly,

$$\begin{aligned}
\mathbb{P}\big[\mu(G_n, \omega, -\gamma) < (1-\delta) E_n(-\gamma)\big] &\leq \mathbb{P}\big[(\mu(G_n, \omega, -\gamma) - E_n(-\gamma))^2 \geq \delta^2 E_n(-\gamma)^2\big] \\
&\leq \frac{1}{\delta^2 E_n(-\gamma)^2} \mathbb{E}\big[(\mu(G_n, \omega, -\gamma) - E_n(-\gamma))^2\big] \\
&\leq \frac{1}{\delta^2 E_n(-\gamma)^2} \big(\mathbb{E}[\mu(G_n, \omega, -\gamma)^2] - E_n(-\gamma)^2\big) \\
&\leq \eta\delta^{-2}.
\end{aligned}$$

By identical arguments,

$$\mathbb{P}\big[\mu^*(G_0(x,t), \omega, \gamma) \geq (1+\delta) E^*_n(\gamma)\big] \leq \eta\delta^{-2} \quad \text{and} \quad \mathbb{P}\big[\mu^*(G_n, \omega, \gamma) < (1-\delta) E^*_n(\gamma)\big] \leq \eta\delta^{-2}.$$

By a union bound, this implies that

$$\mathbb{P}\big[\text{(5-4) and (5-5) hold for all } (x,t)\in G_n\big]\geq 1-4\eta\delta^{-2}, \tag{5-6}$$

so, by choosing $\eta\leq \frac{1}{4}\delta^2$, this has positive probability. Let $\omega\in\Omega$ be an element of this set, which implies ω satisfies (5-4) and (5-5) for all $(x,t)\in G_n$. Using this particular ω, we next show that there exist constants c, C, and $s\in\mathbb{N}$ which only depend on λ, Λ, and d such that

$$c(E_n(-\gamma)+E_n^*(\gamma)-C\gamma^{d+1})\leq (1+\delta)3^{-2(n-s)(d+1)}(E_n(-\gamma)+E_n^*(\gamma)). \tag{5-7}$$

Consider that, by Theorem 3.3, similar to the proof of Lemma 4.4, there exists $n=n(d,\lambda,\Lambda)$ and $u,u^*\in C(G_n(0,3^{2(n+1)}))$ such that

$$u_t+F(D^2u,x,t,\omega)=-\gamma \quad \text{in } G_n(0,3^{2(n+1)})$$

with

$$\inf_{\partial_p G_n(0,3^{2(n+1)})\cap\{t\leq t_0\}} u(x,t)\geq C3^{2n}E_n(-\gamma)^{1/(d+1)} \quad\text{and}\quad \inf_{G_0(0,3^{2(n+1)})} u=\inf_{G_n(0,3^{2(n+1)})} u=0.$$

Similarly, u^* satisfies

$$u_t^*+F(D^2u^*,x,t,\omega^*)=-\gamma \quad \text{in } G_n(0,3^{2(n+1)}),$$

with

$$\inf_{\partial_p G_n(0,3^{2(n+1)})\cap\{t\leq t_0^*\}} u^*(x,t)\geq C3^{2n}E_n^*(\gamma)^{1/(d+1)} \quad\text{and}\quad \inf_{G_0(0,3^{2(n+1)})} u^*=\inf_{G_n(0,3^{2(n+1)})} u^*=0.$$

Let $\tilde{t}=\min\{t_0,t_0^*\}$. We note that the function $u+u^*$ satisfies that

$$u+u^*\geq C3^{2n}(E_n(-\gamma)^{1/(d+1)}+E_n^*(\gamma)^{1/(d+1)}) \quad \text{on } \partial_p G_n(0,3^{2(n+1)})\cap\{t\leq\tilde{t}\}$$

and

$$(u+u^*)_t+\mathcal{M}^+(D^2(u+u^*))\geq -2\gamma \quad \text{in } G_n(0,3^{2(n+1)}).$$

By the Alexandrov–Backelman–Pucci–Krylov–Tso estimate [Wang 1992; Imbert and Silvestre 2012], this implies that

$$u+u^*\geq c3^{2n}[E_n(-\gamma)^{1/(d+1)}+E_n^*(\gamma)^{1/(d+1)}]-C3^{2n}\gamma \quad \text{in } G_n(0,3^{2(n+1)})\cap\{t\leq\tilde{t}\}. \tag{5-8}$$

Next, consider the solutions w, $\tilde{w}$ solving

$$\begin{cases} w_t+F(D^2w,x,t,\omega)=-\gamma & \text{in } G_s(0,3^{2(n+1)}),\\ w=0 & \text{on } \partial_p G_s(0,3^{2(n+1)}),\end{cases}$$

and

$$\begin{cases} w_t^*+F(D^2w^*,x,t,\omega^*)=-\gamma & \text{in } G_s(0,3^{2(n+1)}),\\ w^*=0 & \text{on } \partial_p G_s(0,3^{2(n+1)}),\end{cases}$$

with s, to be chosen, such that $s\leq n$.

We have that

$$w+w^*=0 \quad \text{on } \partial_p G_s(0,3^{2(n+1)})$$

and

$$(w+w^*)_t + \mathcal{M}^-(D^2(w+w^*)) \le -2\gamma \le 0 \quad \text{in } G_s(0, 3^{2(n+1)}).$$

This implies that

$$w+w^* \le 0 \quad \text{in } G_s(0, 3^{2(n+1)}). \tag{5-9}$$

Combining (5-8) and (5-9), we have that, for all $(x,t) \in G_s(0, 3^{2(n+1)}) \cap \{t \le \bar{t}\}$,

$$w(x,t) - u(x,t) + w^*(x,t) - u^*(x,t) \le C3^{2n}\gamma - c3^{2n}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)}). \tag{5-10}$$

Notice that

$$w-u \le 0 \quad \text{on } \partial_p G_s(0, 3^{2(n+1)})$$

and, in $G_s(0, 3^{2(n+1)})$,

$$(w-u)_t + \mathcal{M}^+(D^2(w-u)) \ge 0 \ge (w-u)_t + \mathcal{M}^-(D^2(w-u)).$$

This implies that $w-u \le 0$ in $G_s(0, 3^{2(n+1)})$. Consider the Harnack inequality (Theorem 4.3) applied to $u-w \ge 0$. By the Harnack inequality, rescaled in $G_s(0, 3^{2(n+1)})$ (where $\tilde{Q}$ corresponds to the rescaled $\tilde{Q}$),

$$\sup_{\tilde{Q}}(u-w) \le C \inf_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})}(u-w).$$

This implies that

$$-\sup_{\tilde{Q}}(u-w) \ge -C \inf_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})}(u-w),$$

which yields

$$\inf_{\tilde{Q}}(w-u) \ge C \sup_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})}(w-u). \tag{5-11}$$

Choose s so that $G_0(0, 3^{2(m+1)}) \subseteq Q_{\rho 2^{3^s}}(0, 3^{2(m+1)})$. Since (5-10) holds for all

$$(x,t) \in G_s(0, 3^{2(n+1)}) \cap \{t \le \tilde{t}\} \quad \text{and} \quad \tilde{Q} \subseteq G_s(0, 3^{2(n+1)}) \cap \{t \le \tilde{t}\},$$

we may assume without loss of generality that

$$\inf_{\tilde{Q}}(w-u) \le \tfrac{1}{2}\big(C3^{2n}\gamma - c3^{2n}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)})\big).$$

(If not, then we repeat this analysis for $w^* - u^*$.) By (5-11), this implies that, in $Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})$,

$$w-u \le C\big(3^{2n}\gamma - c3^{2n}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)})\big).$$

In particular, we have that

$$\inf_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})} w \le \inf_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})} u + c\big(C3^{2n}\gamma - 3^{2n}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)})\big).$$

Since $(x_0, t_0) \in G_0(0, 3^{2(n+1)}) \subseteq Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})$, this implies that

$$\inf_{Q_{\rho 2^{3^s}}(0, 3^{2(n+1)})} u = 0,$$

which yields

$$\inf_{Q_{\rho 2 3^s}(0,3^{2(n+1)})} w \le c\big(C3^{2n}\gamma - 3^{2n}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)})\big). \tag{5-12}$$

By Lemma 2.1, since $w = 0$ on $\partial_p G_s(0, 3^{2(n+1)})$,

$$\begin{aligned} 0 &\le \inf_{G_s(0,3^{2(n+1)})} w + c_1 3^{2s}\mu(G_s(0, 3^{2(n+1)}), \omega, -\gamma)^{1/(d+1)} \\ &\le \inf_{Q_{\rho 2 3^s}(0,3^{2(n+1)})} w + c_1 3^{2s}\mu(G_s(0, 3^{2(n+1)}), \omega, -\gamma)^{1/(d+1)}. \end{aligned}$$

By (5-12), this implies

$$\begin{aligned} c3^{2(n-s)(d+1)}(E_n(-\gamma)^{1/(d+1)} + E_n^*(\gamma)^{1/(d+1)} - C\gamma)^{d+1} &\le \mu(G_s(0, 3^{2(n+1)}), \omega, -\gamma) \\ &\le \fint_{G_s(0,3^{2(n+1)})} \mu(G_0(x,t), \omega)\,dx\,dt \\ &\le (1+\delta)E_n(-\gamma) \\ &\le (1+\delta)(E_n(-\gamma) + E_n^*(\gamma)). \end{aligned}$$

This yields

$$3^{2(n-s)(d+1)}c(E_n(-\gamma) + E_n^*(\gamma) - C\gamma^{d+1}) \le (1+\delta)(E_n(-\gamma) + E_n^*(\gamma)),$$

which is equivalent to (5-7).

To conclude, we just need to choose δ, η, and show there is an n sufficiently large to obtain (5-3). Rearranging yields

$$(1 - 3^{-2(n-s)(d+1)} - \delta 3^{-2(n-s)(d+1)})(E_n(-\gamma) + E_n^*(\gamma)) \le C\gamma^{d+1}.$$

Choosing $\delta := 3^{-2s(d+1)}$ and $\eta \le \frac{1}{4}3^{-4s(d+1)}$ yields a choice of $\omega \in \Omega$ such that (5-4) and (5-5) hold, and

$$(1 - 3^{-2(n-s)(d+1)} - 3^{-2n(d+1)})(E_n(-\gamma) + E_n^*(\gamma)) \le C\gamma^{d+1}.$$

For any $n \ge 2s$, we have that

$$E_n(-\gamma) + E_n^*(\gamma) \le C(1 - 3^{-2s(d+1)} - 3^{-4s(d+1)})^{-1}\gamma^{d+1} = C\gamma^{d+1}.$$

This implies that

$$J_n(-\gamma) + J_n^*(\gamma) \le (1+\eta)(E_n(-\gamma)^2 + E_n^*(\gamma)^2) \le C\gamma^{2(d+1)},$$

as asserted. □

We next show how the finite range of dependence assumption (F1) yields a relation between $J_{m+n}(\ell)$ and $J_m(\ell)$ for $n > 0$.

Lemma 5.2. *There exists $c_7 = c_7(d)$ such that, for any ℓ and any $m, n \ge 0$,*

$$J_{m+n}(\ell) \le E_m^2 + \frac{c_7}{3^{n(d+2)}} J_m(\ell). \tag{5-13}$$

Similarly,

$$J^*_{m+n}(-\ell) \le E^{*2}_m + \frac{c_7}{3^{n(d+2)}} J^*_m(-\ell). \tag{5-14}$$

Proof. Since ℓ plays no role, we suppress its dependence. Consider that $G_{m+n} = \bigcup_{i=1}^{3^{n(d+2)}} G^i_m$ for some choice of enumeration of cubes $\{G^i_m\}$. Therefore, for each $u \in S(G_{m+n}, \omega)$,

$$\begin{aligned}
&|\mathcal{P}(G_{m+n}; u)|^2 \\
&\quad= \biggl(\sum_{i=1}^{3^{n(d+2)}} |\mathcal{P}(G^i_m; u)|\biggr)^2 \\
&\quad= \sum_i |\mathcal{P}(G^i_m; u)|^2 + \sum_i \sum_{j \ne i} |\mathcal{P}(G^i_m; u)|\,|\mathcal{P}(G^j_m; u)| \\
&\quad= \sum_{i=1}^{3^{n(d+2)}} |\mathcal{P}(G^i_m; u)|^2 + \sum_{i=1}^{3^{n(d+2)}} \Biggl[\sum_{d[G^i_m, G^j_m] > 1} |\mathcal{P}(G^i_m; u)|\,|\mathcal{P}(G^j_m; u)| + \sum_{d[G^i_m, G^j_m] \le 1} |\mathcal{P}(G^i_m; u)|\,|\mathcal{P}(G^j_m; u)|\Biggr].
\end{aligned}$$

This implies that

$$\begin{aligned}
\mu(G_{m+n}, \omega)^2 \le{}& \frac{1}{3^{2n(d+2)}} \sum_{i=1}^{3^{n(d+2)}} (\mu(G^i_m, u))^2 \\
&+ \frac{1}{3^{2n(d+2)}} \sum_{i=1}^{3^{n(d+2)}} \Biggl[\sum_{d[G^i_m, G^j_m] > 1} \mu(G^i_m, \omega)\mu(G^j_m, \omega) + \sum_{d[G^i_m, G^j_m] \le 1} \mu(G^i_m, \omega)\mu(G^j_m, \omega)\Biggr].
\end{aligned}$$

For each i fixed, if $d[G^i_m, G^j_m] > 1$, then, by (1-7), stationarity, and Lemma 2.8,

$$\mathbb{E}[\mu(G^i_m, \omega)\mu(G^j_m, \omega)] = E^2_m.$$

If $d[G^i_m, G^j_m] \le 1$, then, by the Cauchy–Schwarz inequality and stationarity,

$$\mathbb{E}[\mu(G^i_m, \omega)\mu(G^j_m, \omega)] \le \mathbb{E}[\mu(G_m, \omega)^2] = J_m.$$

For any fixed i, the number of cubes such that $d[G^i_m, G^j_m] \le 1$ is at most 3^{d+1}. Therefore, after taking expectation of both sides, summing over $i = 1, \dots, 3^{n(d+2)}$ copies, this yields that

$$J_{m+n} \le \frac{1}{3^{n(d+2)}} (J_m + (3^{n(d+2)} - 3^{d+1}) E^2_m + 3^{d+1} J_m) \le E^2_m + \frac{C}{3^{n(d+2)}} J_m. \qquad \square$$

Our next lemma shows that, by perturbing ℓ, we can make E and E^* positive.

Lemma 5.3. *Let ℓ be such that*

$$E(\ell) = \lim_{n\to\infty} \mathbb{E}[\mu(G_n, \omega, \ell)] = \lim_{n\to\infty} \mathbb{E}[\mu^*(G_n, \omega, \ell)] = E^*(\ell).$$

There exists $c_8 = c_8(d, \lambda, \Lambda)$ such that, for any $\gamma > 0$ and any n,

$$\mathbb{E}[\mu(G_n, \omega, \ell - \gamma)] \ge c_8 \gamma^{d+1}. \tag{5-15}$$

Analogously,

$$\mathbb{E}[\mu^*(G_n, \omega, -\ell+\gamma)] = \mathbb{E}[\mu(G_n, \omega^*, \ell-\gamma)] \geq c_8\gamma^{d+1}. \tag{5-16}$$

Proof. First, we observe that, by Lemma 4.4, $E(\ell) = 0$. By the subadditive ergodic theorem, we choose $N = N(\delta)$ sufficiently large so that $\mathbb{E}[\mu(G_N, \omega, \ell)] \leq \delta$.

Let w solve

$$\begin{cases} w_t + F(D^2 w, x, t, \omega) = \ell & \text{in } G_N, \\ w = 0 & \text{on } \partial_p G_N. \end{cases}$$

Since $w \in S(G_N, \omega, \ell)$, by Lemma 2.1 we have

$$0 \leq \inf_{G_N} w + c_1 3^{2N} \mu(G_N, \omega, \ell)^{1/(d+1)},$$

which implies that

$$\mathbb{P}\big[w \leq -2^{1/(d+1)} c_1 3^{2N} \delta^{1/(d+1)}\big] \leq \mathbb{P}\big[\mu(G_N, \omega, \ell) \geq 2\delta\big] \leq \tfrac{1}{2}. \tag{5-17}$$

Let $\tilde{w} := w - C\gamma\big(\frac{1}{2}|x|^2 - 3^{2N}\big) + \frac{1}{2}\gamma(3^{2N} - t)$ for C to be chosen. By (5-17),

$$\mathbb{P}\big[\tilde{w} \geq -2c_1 3^{2N}\delta^{1/(d+1)} + C\gamma 3^{2N}\big] \geq \tfrac{1}{2}.$$

Next we consider that there exists $C = C(d, \lambda)$ such that $\tilde{w} \in S(G_N, \omega, \ell-\gamma)$. We verify that

$$\begin{aligned} \tilde{w}_t + F(D^2\tilde{w}, x, t, \omega) &= w_t - \tfrac{1}{2}\gamma + F(D^2 w - C\gamma\,\mathrm{Id}, x, t, \omega) \\ &\geq w_t - \tfrac{1}{2}\gamma + F(D^2 w, x, t, \omega) + \lambda|C\gamma\,\mathrm{Id}\,| \\ &= \ell - \tfrac{1}{2}\gamma + C\lambda\gamma d \geq \ell - \gamma \end{aligned}$$

for $C = C(\lambda, d)$. Since $\tilde{w} \geq 0$ on $\partial_p G_N$, by Lemma 2.1 we have

$$\mathbb{P}\big[\mu(G_N, \omega, \ell-\gamma) \geq C\gamma^{d+1} - C\delta\big] \geq \tfrac{1}{2}.$$

Therefore, for all $n \leq N$,

$$\mathbb{E}[\mu(G_n, \omega, \ell+\gamma)] \geq C(\gamma^{d+1} - \delta).$$

Sending $\delta \to 0$, $N(\delta) \to \infty$, and we have the claim by letting $c_8 = C$. □

We are now ready to obtain a rate of decay on the second moments of μ.

Theorem 5.4. *There exists $\tau = \tau(\lambda, \Lambda, d) \in (0, 1)$ and $c_9 = c_9(\lambda, \lambda, d)$ such that, for all $m \in \mathbb{N}$ and each $M \in \mathbb{S}^d$,*

$$J_m(\bar{F}(M), M) + J_m^*(-\bar{F}(M), M) \leq c_9(1+|M|)^{2(d+1)} K_0^{2(d+1)} \tau^m. \tag{5-18}$$

Proof. We fix $M \in \mathbb{S}^d$ and drop the dependence on $\bar{F}(M)$ (although we mention where it is used). In order to prove (5-18), it is enough to prove that there exists an increasing sequence of integers $\{m_k\}$ such that $|m_{k+1} - m_k| \leq C = C(d, \lambda, \Lambda)$ with

$$J_{m_k}(-3^{-k}) + J_{m_k}^*(3^{-k}) \leq C(1+|M|)^{2(d+1)} K_0^{2(d+1)} 3^{-2k(d+1)}. \tag{5-19}$$

Recall that $|\overline{F}(M)| \le CK_0^{d+1}(1+|M|)^{d+1}$. By (2-8) and scaling, it is enough to assume that we work with

$$J_k := \frac{J_k}{C(1+|M|)^{2(d+1)}K_0^{2(d+1)}},$$

so that $|J_k| \le 1$ and then to prove

$$J_{m_k}(-3^{-k}) + J^*_{m_k}(3^{-k}) \le C3^{-2k(d+1)}. \tag{5-20}$$

Let $m_0 = 0$. Suppose that (5-20) holds for the level m_{k-1}. We would like to find m_k satisfying (5-20) such that $m_k - m_{k-1} \le C$. We aim to set up Lemma 5.1, and then choose $\gamma = 3^{-k}$. Given n_0 and η_0 as in Lemma 5.1, we seek m satisfying (5-13).

Consider that, by Lemma 5.2,

$$J_{m-n_0}(-3^{-k}) \le E^2_{m-n_1}(-3^{-k}) + \frac{c_7}{3^{(n_1-n_0)(d+2)}} J_{m-n_1}(-3^{-k}). \tag{5-21}$$

If we can find a choice of m such that, for a fixed n_1 and η_1,

$$E_{m-n_1}(-3^{-k}) \le (1+\eta_1)^{1/2} E_m(-3^{-k}), \quad E^*_{m-n_1}(3^{-k}) \le (1+\eta_1)^{1/2} E^*_m(3^{-k}), \tag{5-22}$$

and

$$J_{m-n_1}(-3^{-k}) \le (1+\eta_1) J_m(-3^{-k}), \quad J^*_{m-n_1}(3^{-k}) \le (1+\eta_1) J^*_m(3^{-k}), \tag{5-23}$$

then, substituting this into (5-21),

$$\begin{aligned} J_{m-n_0}(-3^{-k}) &\le (1+\eta_1)\left[E^2_m(-3^{-k}) + \frac{c_7}{3^{(n_1-n_0)(d+2)}} J_m(-3^{-k})\right] \\ &\le (1+\eta_1)\left[E^2_m(-3^{-k}) + \frac{c_7}{3^{(n_1-n_0)(d+2)}} J_{m-n_0}(-3^{-k})\right], \end{aligned}$$

which implies that

$$\left[1-(1+\eta_1)\frac{c_7}{3^{(n_1-n_0)(d+2)}}\right] J_{m-n_0}(-3^{-k}) \le (1+\eta_1) E^2_m(-3^{-k}).$$

Similarly, by (5-14),

$$\left[1-(1+\eta_1)\frac{c_7}{3^{(n_1-n_0)(d+2)}}\right] J^*_{m-n_0}(3^{-k}) \le (1+\eta_1) E^{*2}_m(3^{-k}).$$

Choosing $n_1(d,\lambda,\Lambda)$, $\eta_1(d,\lambda,\Lambda)$ so that

$$\left[1-(1+\eta_1)\frac{c_7}{3^{(n_1-n_0)(d+2)}}\right]^{-1}(1+\eta_1) \le 1+\eta_0, \tag{5-24}$$

we may apply Lemma 5.1, to conclude that, for m satisfying (5-22) and (5-23),

$$J_m(-3^{-k}) + J^*_m(3^{-k}) \le C3^{-2k(d+1)}.$$

The problem reduces to finding a choice of m satisfying (5-22) and (5-23) such that m is a bounded distance away from m_{k-1}. This is where we will use the inductive hypothesis. We claim that, for given n_1 and η_1, there exists m such that (5-22) and (5-23) hold, and

$$n_1 \le m \le m_{k-1} + C\log\bigl[C(J_{m_{k-1}}(-3^{-(k-1)}) + J^*_{m_{k-1}}(3^{-(k-1)}))\bigr]. \tag{5-25}$$

Consider that, for all m, by Lemma 5.3, since we are solving with right-hand side $\bar{F}(M)$ (and here is the only place where we use that the right-hand side is $\bar{F}(M)$),

$$c_8 3^{-(k-1)(d+1)} \le E_m(-3^{-(k-1)}) \quad \text{and} \quad c_8 3^{-(k-1)(d+1)} \le E^*_m(3^{-(k-1)}).$$

This implies that, for any N,

$$\prod_{j=1}^{N} \frac{J_{m_{k-1}+(j-1)n_1}(-3^{-(k-1)})}{J_{m_{k-1}+jn_1}(-3^{-(k-1)})} \le C\frac{J_{m_{k-1}}(-3^{-(k-1)})}{3^{-2(k-1)(d+1)}},$$

$$\prod_{j=1}^{N} \frac{J^*_{m_{k-1}+(j-1)n_1}(3^{-(k-1)})}{J^*_{m_{k-1}+jn_1}(3^{-(k-1)})} \le C\frac{J^*_{m_{k-1}}(3^{-(k-1)})}{3^{-2(k-1)(d+1)}},$$

$$\prod_{j=1}^{N} \frac{E_{m_{k-1}+(j-1)n_1}(-3^{-(k-1)})}{E_{m_{k-1}+jn_1}(-3^{-(k-1)})} \le C\frac{E_{m_{k-1}}(-3^{-(k-1)})}{3^{-(k-1)(d+1)}},$$

$$\prod_{j=1}^{N} \frac{E^*_{m_{k-1}+(j-1)n_1}(3^{-(k-1)})}{E^*_{m_{k-1}+jn_1}(3^{-(k-1)})} \le C\frac{E^*_{m_{k-1}}(3^{-(k-1)})}{3^{-(k-1)(d+1)}}.$$

Since each individual term in the product is bounded from below by 1, this implies that there exists some element j^i for $i = 1, 2, 3, 4$ such that

$$\frac{J_{m_{k-1}+(j^1-1)n_1}(-3^{-(k-1)})}{J_{m_{k-1}+j^1n_1}(-3^{-(k-1)})} \le C\left(\frac{J_{m_{k-1}}(-3^{-(k-1)})}{3^{-2(k-1)(d+1)}}\right)^{\frac{1}{N}},$$

$$\frac{J^*_{m_{k-1}+(j^2-1)n_1}(3^{-(k-1)})}{J^*_{m_{k-1}+j^2n_1}(3^{-(k-1)})} \le C\left(\frac{J^*_{m_{k-1}}(3^{-(k-1)})}{3^{-2(k-1)(d+1)}}\right)^{\frac{1}{N}},$$

$$\frac{E_{m_{k-1}+(j^3-1)n_1}(-3^{-(k-1)})}{E_{m_{k-1}+j^3n_1}(-3^{-(k-1)})} \le C\left(\frac{J_{m_{k-1}}(-3^{-(k-1)})}{3^{-2(k-1)(d+1)}}\right)^{\frac{1}{2N}},$$

$$\frac{E^*_{m_{k-1}+(j^4-1)n_1}(3^{-(k-1)})}{E^*_{m_{k-1}+j^4n_1}(3^{-(k-1)})} \le C\left(\frac{J^*_{m_{k-1}}(3^{-(k-1)})}{3^{-2(k-1)(d+1)}}\right)^{\frac{1}{2N}}.$$

Let

$$N := \left\lceil C\frac{\log\bigl[3^{2(k-1)(d+1)}(J_{m_{k-1}}(-3^{-(k-1)}) + J^*_{m_{k-1}}(3^{k-1}))\bigr]}{\log(1+\delta_1)}\right\rceil$$

and set $m := m_{k-1} + jn_1$ for $j := \max_i\{j^i\} \le N$. Applying the monotonicity, this choice of m satisfies (5-22) and (5-23). Define $m_k := m$, and this implies, by the inductive hypothesis, that

$$\begin{aligned} m_k &\le m_{k-1} + C \log\big[3^{2(k-1)(d+1)}(J_{m_{k-1}}(-3^{-(k-1)}) + J^*_{m_{k-1}}(3^{k-1}))\big] \\ &\le m_{k-1} + C \log[C3^{2(k-1)(d+1)}3^{-2(k-1)(d+1)}] \le m_{k-1} + C. \end{aligned}$$

This completes the induction and the proof of (5-19). By the monotonicity in the right-hand side ℓ, this actually yields a sequence $\{m_k\}$ such that $|m_k - m_{k-1}| \le C$ for all k and

$$J_{m_k} + J^*_{m_k} \le C3^{-2k(d+1)}.$$

Using the monotonicity of J_m in m to interpolate between points $m = 3^{m_k}$, we obtain (5-18) for some c_9. □

Using this rate on the decay of the second moments, we apply Chebyshev's inequality to obtain a rate on the decay of μ.

Corollary 5.5. *For every $p < d+2$, there exists $c = c(p, \lambda, \Lambda, d)$ and $\alpha = \alpha(\lambda, \Lambda, p, d)$ such that, for all $m \in \mathbb{N}$ and all $\nu \ge 1$,*

$$\mathbb{P}\big[\mu(G_m, \omega, \bar{F}(M), M) \ge (1+|M|)^{d+1}K_0^{d+1}3^{-m\alpha}\nu\big] \le \exp(-c\nu 3^{mp}) \tag{5-26}$$

and

$$\mathbb{P}\big[\mu^*(G_m, \omega, \bar{F}(M), M) \ge (1+|M|)^{d+1}K_0^{d+1}3^{-m\alpha}\nu\big] \le \exp(-c\nu 3^{mp}). \tag{5-27}$$

Proof. We only prove (5-26), since (5-27) follows by identical arguments. Without loss of generality, we assume that $M = 0$ and we drop the dependence on $\bar{F}(0)$.

Fix $m \in \mathbb{N}$ and let $n \in \mathbb{N}$ to be chosen. We consider decomposing $G_{m+n+1} = \bigcup_{i=1}^{3^{d+2}} \mathcal{G}_n^i$, where $\mathcal{G}_n^i = \bigcup_{j=1}^{3^{m(d+2)}} G_n^{ij}$ is a collection of subcubes of size G_n such that each of the subcubes of size G_n is separated by distance at least 1.

By the finite range of dependence assumption (F1), for each i,

$$\mu(G_n^{ij}, \omega) \;\text{ and }\; \mu(G_n^{ik}, \omega) \;\text{ are independent if }\; j \ne k. \tag{5-28}$$

Using this decomposition yields that

$$\begin{aligned} \log \mathbb{E}\big[\exp(\nu 3^{m(d+2)}\mu(G_{m+n+1}, \omega))\big] &\le \log \mathbb{E}\Bigg[\prod_{i=1}^{3^{d+2}} \prod_{j=1}^{3^{m(d+2)}} \exp(\nu 3^{-(d+2)}\mu(G_n^{ij}, \omega))\Bigg] \\ &\le 3^{-(d+2)} \sum_{i=1}^{3^{(d+2)}} \log \mathbb{E}\Bigg[\prod_{j=1}^{3^{m(d+2)}} \exp(\nu\mu(G_n^{ij}, \omega))\Bigg] \\ &= 3^{-(d+2)} \sum_{i=1}^{3^{(d+2)}} \log\Bigg(\prod_{j=1}^{3^{m(d+2)}} \mathbb{E}\big[\exp(\nu\mu(G_n^{ij}, \omega))\big]\Bigg) \\ &= 3^{m(d+2)} \log \mathbb{E}\big[\exp(\nu\mu(G_n, \omega))\big], \end{aligned}$$

where the last line holds by stationarity. Moreover, if we choose $\nu = CK_0^{-1/(d+1)}$, then $\nu\mu(G_n, \omega) \le 1$ almost surely. Using the elementary inequalities

$$\begin{cases} \exp(s) \le 1 + 2s & \text{for all } 0 \le s \le 1, \\ \log(1+s) \le s & \text{for all } s \ge 0, \end{cases}$$

yields that, for this choice of ν,

$$\log \mathbb{E}\big[\exp(CK_0^{-(d+1)} 3^{m(d+2)} \mu(G_{m+n+1}, \omega))\big] \le 3^{m(d+2)} \mathbb{E}[CK_0^{-(d+1)} \mu(G_n, \omega)] \le C 3^{m(d+2)} \tau^n \tag{5-29}$$

by Theorem 5.4.

Therefore, by Chebyshev's inequality and (5-29), this yields that

$$\begin{aligned} \mathbb{P}\big[\mu(G_{m+n+1}, \omega) \ge K_0^{d+1} \nu\big] &\le \mathbb{P}\big[\exp(K_0^{-(d+1)} 3^{m(d+2)} \mu(G_{m+n+1}, \omega)) \ge \exp(3^{m(d+2)} \nu)\big] \\ &\le C \exp(-3^{m(d+2)}(\nu - \tau^n)). \end{aligned}$$

Letting $\nu = \frac{1}{2}\tau^n \nu$ and using that $\nu \ge 1$, we have that

$$\mathbb{P}\big[\mu(G_{m+n+1}, \omega) \ge C\tau^n K_0^{d+1} \nu\big] \le C \exp(-3^{m(d+2)} \tau^n \nu).$$

Choosing $n \sim \lfloor (mp \log 3)/(2(p \log 3 + |\log \tau|)) \rfloor \le \frac{1}{2}m$ implies that $c3^{-mp} \le \tau^n \le C3^{-mp}$, which yields that

$$\mathbb{P}\big[\mu(G_{m+n+1}, \omega) \ge C3^{-mp} K_0^{d+1} \nu\big] \le C \exp(-3^{m(d+2-p)} \nu).$$

Relabeling $m = m + n + 1$ and $p = d + 2 - p$ yields that there exists $\alpha = \alpha(\lambda, \Lambda, p, d)$ such that

$$\mathbb{P}\big[\mu(G_m, \omega) \ge C3^{-m\alpha} K_0^{d+1} \nu\big] \le C \exp(-3^{mp} \nu). \qquad \square$$

6. The proof of Theorem 1.1

We finally present the rate for homogenization in probability using Theorem 5.4. This follows a general procedure which has been shown in [Caffarelli and Souganidis 2010; Armstrong and Smart 2014b; Lin 2015]. However, for completeness we provide the argument here as well, similar to the approach of [Armstrong and Smart 2014b]. As mentioned in the above references, if the limiting function u is $C^2(\mathbb{R}^{d+1})$ (i.e., $C^2(\mathbb{R}^d) \cap C^1([0, T])$), then obtaining a rate for the homogenization is straightforward. Studying $\lim_{\varepsilon \to 0} w^\varepsilon$, where w^ε solves (4-1), is equivalent to the stochastic homogenization of (1-1) when the limiting function is of the form $u(x, t) = bt + \frac{1}{2}x \cdot Mx$. By (4-12) and Chebyshev's inequality, a rate on the decay of $\mu(G_{1/\varepsilon}, \omega)$ immediately yields a rate in probability for the decay of w^ε. If $u \in C^2$, then, by replacing u with its second-order Taylor series expansion with cubic error, we obtain a rate for $u^\varepsilon - u$. In general, since u is not necessarily C^2, we must argue that one can still approximate u by a quadratic expansion. This type of approximation is the motivation for the theory of δ-viscosity solutions, which was introduced in the elliptic setting in [Caffarelli and Souganidis 2010] and generalized to the parabolic setting by Turanova [2015]. The rate in [Lin 2015] was obtained by using this regularization procedure.

For clarity and for a more general approach, we choose to present the argument in terms of a quantified comparison principle as in [Armstrong and Smart 2014b]. We revert to quantifying the traditional "doubling variables" arguments used in the theory of viscosity solutions (see for example [Crandall et al. 1992; Crandall 1997]). We are informed that this is related to forthcoming work by Armstong and Daniel [2015], who generalize this method to finite difference schemes for fully nonlinear, uniformly parabolic equations. The next series of results are entirely deterministic and therefore we suppress the dependence on the random parameter ω.

We first present a result relating the measure of the parabolic subdifferential to the measure of the corresponding touching points in physical space-time.

Proposition 6.1. *Let u and v be such that*

$$u_t + \mathcal{M}^-(D^2u) - R_0 \le 0 \le v_t + \mathcal{M}^+(D^2v) + R_0 \quad \text{in } U_T. \tag{6-1}$$

Assume $\delta > 0$ and let $V = \overline{V} \subseteq U_T \times U_T$ and $W \subseteq \mathbb{R}^{d+1} \times \mathbb{R}^{d+1}$ be such that, for all $((p,h),(q,k)) \in W$,

$$\begin{aligned}\Big\{(x,t,y,s) : \sup_{U_T\times U_T:\tau\le t,\sigma\le s} & u(\xi,\tau) - v(\eta,\sigma) - \frac{1}{2\delta}[|\xi-\eta|^2 + (\tau-\sigma)^2] - p\cdot\xi - q\cdot\eta \\ &= u(x,t) - v(y,s) - \frac{1}{2\delta}[|x-y|^2+(t-s)^2] - p\cdot x - q\cdot y, \\ h = u(x,t) - \frac{1}{2\delta}[|x-y|^2+(t-s)^2] - p\cdot x,\ & k = -v(y,s) - \frac{1}{2\delta}[|x-y|^2+(t-s)^2] - q\cdot y\Big\} \subseteq V.\end{aligned}$$

Then there exists a constant $C = C(\lambda, \Lambda, d, U_T)$ such that

$$|W| \le C(R_0 + \delta^{-1})^{2d+2}|V|. \tag{6-2}$$

Proof. Without loss of generality, we may assume by scaling that $U_T \subseteq Q_1(0,1)$. As usual, we constantly relabel C for a constant which only depends on λ, Λ, and d. For $i = 1, 2$, let $(x_i, t_i, y_i, s_i, p_i, q_i, h_i, k_i)$ satisfy

$$\begin{aligned}\sup_{U_T\times U_T,\tau\le t_i,\sigma\le s_i} u(x,\tau) - v(y,\sigma) - \frac{1}{2\delta}(|x-y|^2+(\tau-\sigma)^2) - p_i\cdot x - q_i\cdot y \\ = u(x_i,t_i) - v(y_i,s_i) - \frac{1}{2\delta}(|x_i-y_i|^2+(t_i-s_i)^2) - p_i\cdot x_i - q_i\cdot y_i \\ = h_i + k_i,\end{aligned}$$

and let

$$\Delta = (|x_1-x_2|^2 + |y_1-y_2|^2 + |t_1-t_2| + |s_1-s_2|)^{1/2}. \tag{6-3}$$

We claim that

$$(|p_1-p_2|^2 + |q_1-q_2|^2 + |h_1-h_2|^2 + |k_1-k_2|^2)^{1/2} \le C(1+\delta^{-1})\Delta + o(\Delta) \tag{6-4}$$

as $|\Delta| \to 0$.

If (6-4) holds, then one can obtain (6-2) using standard measure-theoretic arguments. A priori, this may not be apparent since the left-hand side of (6-4) corresponds to the Euclidean distance between points in $\mathbb{R}^{d+1}$, whereas Δ corresponds to the parabolic distance under the metric $d[\cdot,\cdot]$. However, the parabolic cylinders have the appropriate doubling property with respect to Lebesgue measure, and thus standard measure-theoretic arguments apply.

We prove a series of claims, using standard techniques in the method of doubling variables.

Claim. *For each i,*

$$|t_i - s_i| \le \delta R_0 + C. \tag{6-5}$$

Consider that the map

$$(x,t) \to u(x,t) - \frac{1}{2\delta}[|x-y_1|^2 + (t-s_1)^2] - p_1 \cdot x$$

achieves its maximum over $U \times (0, t_1]$ at (x_1, t_1). Therefore, by (6-1),

$$\frac{1}{\delta}(t_1 - s_1) + \mathcal{M}^-(\delta^{-1}\,\mathrm{Id}) \le R_0,$$

implying that

$$t_1 - s_1 \le \delta[R_0 - (-C\delta^{-1})] = \delta R_0 + C. \tag{6-6}$$

Similarly, the map

$$(y,s) \to v(y,s) + \frac{1}{2\delta}[|x_1-y|^2 + (t_1-s)^2] + q_1 \cdot y$$

achieves its minimum over $U \times (0, s_1]$ at (y_1, s_1). By (6-1),

$$t_1 - s_1 \ge \delta(-R_0 - C\delta^{-1}) = -\delta R_0 - C. \tag{6-7}$$

Combining (6-6) and (6-7) yields (6-5).

Claim. *Let $u_t + \mathcal{M}^+(D^2u) \ge -1$ in Q_1. Let $(p_1, h_1) \in \mathcal{P}((x_1, t_1); u)$ and $(p_2, h_2) \in \mathcal{P}((x_2, t_2); u)$. Then*

$$|p_1 - p_2|^2 + |h_1 - h_2|^2 \le C(|x_1 - x_2|^2 + |t_1 - t_2|^2 + |x_1 - x_2|^4 + |t_1 - t_2|). \tag{6-8}$$

Without loss of generality, by subtracting a plane and translating, we may assume that $(p_2, h_2) = (0, 0)$ and $(x_2, t_2) = (0, 0)$. The claim will follow from the regularity of Γ^u (Lemma 2.3). Since $(x_1, t_1), (0, 0) \in \{u = \Gamma^u\}$ and $D\Gamma^u$ is Lipschitz continuous, this implies that

$$|p_1| \le C(|x_1|^2 + |t_1|)^{1/2}.$$

To estimate $|h_1|$, we again apply the regularity of Γ^u and the bound on $|p_1|$ to conclude that

$$|h_1| = |h_1 - h_2| = |u(x_1, t_1) - p_1 \cdot x_1 - u(x_2, t_2)| \le C(|x_1|^2 + |t_1|)^{1/2}(1 + |x_1|).$$

Therefore,

$$|h_1|^2 \le C^2(|x_1|^2 + |t_1|)(1 + |x_1|)^2 \le C(|x_1|^2 + |t_1|^2 + |x_1|^4 + |t_1|).$$

Combining these observations yields (6-8).

Next, we apply these observations to the parabolic subdifferentials. For simplicity, we adopt some notation. Without loss of generality, assume that $s_1 \geq s_2$. Let $T_{\min} := \min\{t_1, t_2, s_2\}$ and $T_{\max} := \max\{t_1, t_2, s_1\}$. Notice that, by (6-5), $T_{\max} - T_{\min} \leq \delta R_0 + C + \Delta^2 := \gamma^2$. Therefore, $(x_1, t_1), (x_2, t_2) \in Q_\gamma(x_1, T_{\max})$. Let

$$\tilde{u}(x, t) := -u(x, t) + \frac{1}{2\delta}[|x - y_1|^2 + (t - s_1)^2].$$

This implies that

$$\begin{aligned}
\tilde{u}_t + \mathcal{M}^+(D^2\tilde{u}) &= -u_t + \delta^{-1}(t - s_1) + \mathcal{M}^+(-D^2u + \delta^{-1}\,\mathrm{Id}) \\
&\geq -u_t + \delta^{-1}(t - s_1) - \mathcal{M}^-(D^2u) - \delta^{-1}C \\
&\geq -R_0 - C(1 + \delta R_0 + \Delta^2)\delta^{-1} \\
&\geq -C(R_0 + \delta^{-1}(1 + \Delta^2)) \quad \text{in } Q_\gamma(x_1, T_M).
\end{aligned} \tag{6-9}$$

We next find elements in the parabolic subdifferential of $\tilde{u}$.

Claim. *We have*

$$(-p_1, \tilde{u}(x_1, t_1) + p_1 \cdot x_1) \in \mathcal{P}((x_1, t_1); \tilde{u}). \tag{6-10}$$

Since

$$u(x_1, t_1) - \frac{1}{2\delta}[|x_1 - y_1|^2 + (t_1 - s_1)^2] - p_1 \cdot x_1 \geq u(x, t) - \frac{1}{2\delta}[|x - y_1|^2 + (t - s_1)^2] - p_1 \cdot x$$

for all $t \leq t_1$ and $x \in U$, this implies that

$$\tilde{u}(x_1, t_1) - (-p_1 \cdot x_1) = -u(x_1, t_1) + \frac{1}{2\delta}(|x_1 - y_1|^2 + (t_1 - s_1)^2) + p_1 \cdot x_1 \leq \tilde{u}(x, t) - (-p_1 \cdot x)$$

for all $t \leq t_1$ and $x \in U$. This yields (6-10).

Claim. *We have*

$$\left(-p_2 + \frac{y_2 - y_1}{\delta}, \tilde{u}(x_2, t_2) + \left(p_2 - \frac{y_2 - y_1}{\delta}\right) \cdot x_2\right) \in \mathcal{P}((x_2, t_2); \tilde{u}). \tag{6-11}$$

Since

$$\begin{aligned}
&-u(x, t) + \frac{1}{2\delta}[|x - y_2|^2 + (t - s_2)^2] + p_2 \cdot x \\
&\qquad = \tilde{u}(x, t) + \frac{1}{2\delta}[|x - y_2|^2 + (t - s_2)^2 - |x - y_1|^2 - (t - s_1)^2] + p_2 \cdot x \\
&\qquad = \tilde{u}(x, t) + \left(\frac{1}{\delta}(-y_2 + y_1) + p_2\right) \cdot x + \frac{1}{2\delta}[(t - s_2)^2 - (t - s_1)^2 + |y_2^2| - |y_1|^2],
\end{aligned}$$

we obtain that

$$\begin{aligned}
&\tilde{u}(x_2, t_2) + \left(\frac{1}{\delta}(-y_2 + y_1) + p_2\right) \cdot x_2 + \frac{1}{2\delta}[(t_2 - s_2)^2 - (t_2 - s_1)^2] \\
&\qquad \leq \tilde{u}(x, t) + \left(\frac{1}{\delta}(-y_2 + y_1) + p_2\right) \cdot x + \frac{1}{2\delta}[(t - s_2)^2 - (t - s_1)^2].
\end{aligned}$$

Simplifying yields that

$$\tilde{u}(x_2, t_2) + \left(\frac{1}{\delta}(-y_2 + y_1) + p_2\right) \cdot x_2 + \frac{1}{\delta}[-(t_2 - t)(s_2 - s_1)] \leq \tilde{u}(x, t) + \left(\frac{1}{\delta}(-y_2 + y_1) + p_2\right) \cdot x.$$

Therefore, for $t \le t_2$, since $s_1 \ge s_2$,

$$\tilde{u}(x_2, t_2) + \Big(\frac{1}{\delta}(-y_2 + y_1) + p_2\Big) \cdot x_2 \le \tilde{u}(x, t) + \Big(\frac{1}{\delta}(-y_2 + y_1) + p_2\Big) \cdot x,$$

which yields the claim.

By combining (6-8), (6-9), (6-10), and (6-11),

$$\begin{aligned}\Big|p_1 - p_2 + \frac{1}{\delta}(y_2 - y_1)\Big|^2 + \Big|\tilde{u}(x_1, t_1) + p_1 \cdot x_1 - \tilde{u}(x_2, t_2) - \Big(p_2 - \frac{1}{\delta}(y_2 - y_1)\Big) \cdot x_2\Big|^2 \\ \le C[R_0 + \delta^{-1}(1 + \Delta^2)]^2(|x_1 - x_2|^2 + |t_1 - t_2|^2 + |x_1 - x_2|^4 + |t_1 - t_2|).\end{aligned}$$

Recall that

$$-\tilde{u}(x_1, t_1) - p_1 \cdot x_1 = h_1$$

and

$$\begin{aligned}-\tilde{u}(x_2, t_2) - \Big(p_2 - \frac{1}{\delta}(y_2 - y_1)\Big) \cdot x_2 &= h_2 + \frac{1}{2\delta}(|y_2|^2 - |y_1|^2) + \frac{1}{2\delta}[(t_2 - s_2)^2 - (t_2 - s_1)^2] \\ &= h_2 + \frac{1}{2\delta}[|y_2|^2 - |y_1|^2 + s_2^2 - s_1^2 - 2t_2(s_2 - s_1)].\end{aligned}$$

Collecting terms yields that

$$\begin{aligned}|p_1 - p_2|^2 + |h_1 - h_2|^2 &\le C[R_0 + \delta^{-1}(1 + \Delta^2)]^2[|x_1 - x_2|^2 + |t_1 - t_2|^2 + |x_1 - x_2|^4 + |t_1 - t_2|] \\ &\qquad + \frac{1}{\delta^2}|y_2 - y_1|^2 + \frac{1}{4\delta^2}[|y_2|^2 - |y_1|^2 + s_2^2 - s_1^2 - 2t_2(s_2 - s_1)]^2 \\ &\le C[R_0 + \delta^{-1}(1 + \Delta^2)]^2\Delta^2 + \frac{1}{\delta^2}o(\Delta^2) \\ &\le C[R_0 + \delta^{-1}]^2\Delta^2 + o(\Delta^2),\end{aligned}$$

which implies that

$$(|p_1 - p_2|^2 + |h_1 - h_2|^2)^{1/2} \le C(R_0 + \delta^{-1})\Delta + o(\Delta).$$

An analogous argument yields that

$$(|q_1 - q_2|^2 + |k_1 - k_2|^2)^{1/2} \le C(R_0 + \delta^{-1})\Delta + o(\Delta).$$

Combined, this yields (6-4). □

Next, we show that, if $|u - u^\varepsilon|$ is large somewhere, then we can find a matrix M^* and a parabolic cube G^* such that $\mu(G^*, \bar{F}(M^*), M^*)$ is very large. We mention that both M^* and G^* come from a countable family of matrices and cubes. In order to select M^* and G^*, we must construct the appropriate approximation of u to argue that u is close to a quadratic expansion. We will employ the $W^{3,\alpha}$ estimate proven in [Daniel 2015], which yields an estimate on the measure of points which can be well-approximated by a quadratic expansion. We state the result slightly differently than it appears in [Daniel 2015], in order to readily apply it for our purposes.

Theorem 6.2 [Daniel 2015, Theorem 1.2]. *Let $u_t + F(D^2u) = 0$ in Q_1, $u = g$ on $\partial_p Q_1$, with F uniformly parabolic. Let $Q \subseteq Q_1$. For each $\kappa > 0$, let*

$$\Sigma_\kappa := \big\{(x,t) \in Q_1 : \exists (M,\xi,b) \in \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R} \text{ with } |M| \le \kappa \text{ such that, for all } (y,s) \in Q_1 \text{ with } s \le t,$$
$$\big|u(y,s) - u(x,t) - b(s-t) - \xi\cdot(y-x) - \tfrac{1}{2}(y-x)\cdot M(y-x)\big| \le \tfrac{1}{6}\kappa(|x-y|^3 + |s-t|^{3/2})\big\}.$$

There exists $C = C(\lambda,\Lambda,d)$ and $\alpha = \alpha(\lambda,\Lambda,d)$ such that, for every $\kappa > 0$,

$$\big|Q_1 \setminus \big(\Sigma_\kappa \cap Q_{1/2}(0,-\tfrac{1}{4})\big)\big| \le C\left(\frac{\kappa}{\sup_{Q_1}\big([|u| + |F(0,\cdot,\cdot)|] + \|g\|_{C^{0,1}(\partial_p Q_1)}\big)}\right)^{-\alpha}.$$

We note that Σ_κ corresponds to the set of points which can be touched monotonically in time by a quadratic expansion with controllable error. Moreover, the points in Σ_κ are touched from above and below by polynomials. We are now ready to show the existence of M^* and G^*. For simplicity, we say that a function $\Phi : U_T \times U_T$ achieves a monotone maximum at (x_0,t_0,y_0,s_0) if $\Phi(x_0,t_0,y_0,s_0) \ge \Phi(x,t,y,s)$ for all $x, y \in U$ and all $t \le t_0$, $s \le s_0$.

Proposition 6.3. *Let u and v satisfy*

$$\begin{cases} u_t + \overline{F}(D^2u) = f(x,t) = v_t + F(D^2v,x,t) & \text{in } U_T,\\ u = v = g(x,t) & \text{on } \partial_p U_T,\end{cases}$$

such that

$$\|\overline{F}(0)\|_{L^\infty(U_T)} + \sup\|F(0,\cdot,\cdot)\|_{L^\infty(U_T)} + \|g\|_{C^{0,1}(\partial_p U_T)} + \|f\|_{C^{0,1}(U_T)} \le R_0 < +\infty.$$

There exists an exponent $\sigma = \sigma(\lambda,\Lambda,d) \in (0,1)$ and constants $c = c(\lambda,\Lambda,d,U_T)$, $C = C(\lambda,\Lambda,d,U_T)$ such that, for any $l \le \eta$, if

$$A := \sup_{U_T}(u-v) \ge CR_0\eta^\sigma > 0, \tag{6-12}$$

then there exists $M^* \in \mathbb{S}^d$, $(y^*,s^*) \in U_T$ such that:

- $|M^*| \le \eta^{\sigma-1}$,
- *$l^{-1}M^*$, $\eta^{-1}y^*$, and $\eta^{-2}s^*$ have integer entries,*
- $\mu((y^*,s^*) + \eta G_0, \overline{F}(M^*), M^*) \ge cA^{d+1}$,

where $\eta G_0 = \big(-\tfrac{1}{2}\eta, \tfrac{1}{2}\eta\big]^d \times (-\eta^2, 0]$.

Proof. As usual, c and C will denote constants which depend on universal quantities, which will vary line by line. We first point out some simplifications which we take without loss of generality. We assume that $R_0 = 1$ and $U_T \subseteq Q_1(0,1)$, and appropriately renormalize.

Next, we claim that we may replace v by $\tilde{v}$ solving

$$\begin{cases} \tilde{v}_t + F(D^2\tilde{v},x,t) = f(x,t) + cA & \text{in } U_T,\\ \tilde{v} = v & \text{on } \partial_p U_T.\end{cases} \tag{6-13}$$

The Alexandrov–Backelman–Pucci–Krylov–Tso estimate [Wang 1992; Imbert and Silvestre 2012] yields that

$$\tilde{v} - v \le CA \quad \text{in } U_T,$$

so, by adjusting the constant in (6-12), we may take the replacement at no cost.

Finally, we point out that, by the Krylov–Safonov estimates [Wang 1992; Imbert and Silvestre 2012], u and v are Hölder continuous and, since $R_0 \le 1$, there exists $\alpha(\lambda, \Lambda, d) \in (0, 1)$ such that

$$\|u\|_{C^{0,\alpha}(\bar{U}_T)} + \|v\|_{C^{0,\alpha}(\bar{U}_T)} \le C. \tag{6-14}$$

Without loss of generality, assume that $\alpha \le \frac{1}{2}$. Since $u = v$ on $\partial_p U_T$, this implies that, for all $(x, t), (y, s) \in U_T$,

$$|u(x, t) - v(y, s)| \le C\big(d[(x, t), \partial_p U_T]^\alpha + d[(y, s), \partial_p U_T]^\alpha + d[(x, t), (y, s)]^\alpha\big).$$

Consider the function

$$\Phi(x, t, y, s, p, q) = u(x, t) - v(y, s) - \frac{1}{2\delta}[|x - y|^2 + (t - s)^2)] - p \cdot x - q \cdot y.$$

Suppose there exists a point (x_0, t_0) such that $u(x_0, t_0) - v(x_0, t_0) \ge \frac{3}{4}A$. This implies that

$$\Phi(x_0, t_0, x_0, t_0, 0, 0) \ge \tfrac{3}{4}A.$$

Let

$$U_T(\rho) := \{(x, t) \in U_T \times U_T : d[(x, t), \partial_p U_T] \ge \rho\}.$$

Let $p, q \in B_r$, where we define $r := \frac{1}{8}A$. We would like to show that $\Phi(\cdot, \cdot, \cdot, \cdot, p, q)$ achieves it monotone maximum in $U_T(\rho) \times U_T(\rho)$ for some choice of ρ.

We note that

$$\begin{aligned}
&\Phi(x, t, y, s, p, q) \\
&\quad = u(x, t) - v(y, s) - \frac{1}{2\delta}[|x - y|^2 + (t - s)^2] - p \cdot x - q \cdot y \\
&\quad \le C\big(d[(x, t), \partial_p U_T]^\alpha + d[(y, s), \partial_p U_T]^\alpha + d[(x, t), (y, s)]^\alpha\big) - \frac{1}{2\delta}[|x - y|^2 + (t - s)^2] + 2r.
\end{aligned}$$

By Young's inequality,

$$|x - y|^\alpha = A^{(2-\alpha)/2}[A^{-(2-\alpha)/\alpha}|x - y|^2]^{\alpha/2} \le \frac{1}{8C}A + CA^{-(2-\alpha)/\alpha}|x - y|^2$$

and

$$|t - s|^{\alpha/2} = A^{(4-\alpha)/4}[A^{-(4-\alpha)/\alpha}|t - s|^2]^{\alpha/4} \le \frac{1}{8C}A + CA^{-(4-\alpha)/\alpha}(t - s)^2.$$

Assume $A \le 1$. This implies that $A^{-(2-\alpha)/\alpha} \le A^{-(4-\alpha)/\alpha}$. Therefore,

$$\begin{aligned}
&\Phi(x, y, t, s, p, q) \\
&\quad \le Cd[(x, t), \partial_p U_T]^\alpha + Cd[(y, s), \partial_p U_T]^\alpha + \tfrac{1}{4}A + \tfrac{1}{4}A + C\Big(A^{-(4-\alpha)/\alpha} - \frac{1}{2\delta}\Big)[|x - y|^2 + (t - s)^2].
\end{aligned}$$

By letting

$$\delta \le \tfrac{1}{2} A^{(4-\alpha)/\alpha}, \tag{6-15}$$

we have that

$$\Phi(x, y, t, s, p, q) \le Cd[(x, t), \partial_p U_T]^\alpha + C[d(y, s), \partial_p U_T]^\alpha + \tfrac{1}{2} A.$$

Therefore, letting $\rho := CA^{1/\alpha}$ yields that, for any $p, q \in B_r$, Φ achieves its monotone maximum in $U_T(\rho) \times U_T(\rho)$.

Using the language of Proposition 6.1, we choose $W \subseteq \mathbb{R}^{d+1}$ such that $Q_r \times Q_r \subseteq W$. This yields that

$$\begin{aligned} V := \big\{ &(x, t, y, s) \in U_T \times U_T : \text{for some } (p, q) \in B_r \times B_r, \\ &\Phi(\cdot, \cdot, \cdot, \cdot, p, q) \text{ achieves its monotone maximum at } (x, t, y, s) \text{ for appropriate } (h, k) \in \mathbb{R}^2 \big\} \\ &\subseteq U_T(\rho) \times U_T(\rho). \end{aligned}$$

By Proposition 6.1, this implies that

$$|V| \ge C(1 + \delta^{-1})^{-2d-2} r^{2d+2} \ge C(1 + A^{-(4-\alpha)/\alpha})^{-2d-2} A^{2d+2} \ge CA^{(8d+8)/\alpha}.$$

If we define the projection $\pi : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ by $\pi((A, B)) = A$, we have that

$$\pi(V) \ge |U_T|^{-1} |V| \ge |Q_1|^{-1} |V| \ge CA^{(8d+8)/\alpha}. \tag{6-16}$$

Finally, we note that, for every $((x, t), (y, s)) \in V$, since $\Phi(x, t, y, s, p, q) \ge 0$ for some $p, q \in B_r \subseteq B_1$, $\alpha \le \frac{1}{2}$, and $A \le 1$, this implies that

$$|x - y|^2 + |t - s|^2 \le C\delta \le CA^{(4-\alpha)/\alpha} \le CA^6. \tag{6-17}$$

Next, we use (6-16) to show that there are points in $\pi(V)$ where u can be approximated by a quadratic expansion. Let Σ_κ as in the $W^{3,\alpha}$ estimate (Theorem 6.2).

By the $W^{3,\alpha}$ estimate, assuming that $U_T \subseteq Q_1$,

$$|U_T \setminus \Sigma_\kappa(U_T)| \le \big| Q_1 \setminus \Sigma_\kappa(U_T) \cap Q_{1/2}\big(0, -\tfrac{1}{4}\big) \big| \le C\kappa^{-\alpha}. \tag{6-18}$$

Although a priori the two α's in (6-16) and (6-18) are not necessarily the same, we can assume without loss of generality they are the same by taking the minimum of the two.

Thus, if we let $\kappa \ge CA^{-4(d+2)/\alpha^2}$, then

$$|U_T \setminus \Sigma_\kappa(U_T)| < |\pi(V)|,$$

which implies that $\pi(V) \cap \Sigma_\kappa \ne \varnothing$. This implies that there are points of $\pi(V)$ where u can be touched monotonically in time by a quadratic expansion with controllable error, and the function Φ achieves it monotone maximum there.

Finally, we show that there exist M^*, y^*, s^*, and G^* which satisfy the conclusion of the proposition. By the previous step, there exists $(x_1, t_1, y_1, s_1) \in V$ with $(x_1, t_1) \in \Sigma_\kappa$. In other words, there exist $p, q \in B_r$ such that

$$\Phi(x_1, t_1, y_1, s_1, p, q) = \sup_{U_T(\rho) \times U_T(\rho), \tau \le t_1, \sigma \le s_1} \Phi(x, \tau, y, \sigma, p, q),$$

and (M, ξ, b) such that $|M| \le \kappa$ and, for all $(x,t) \in U_T$, $t \le t_1$,

$$\left|u(x,t) - u(x_1,t_1) - b(t-t_1) - \xi \cdot (x-x_1) - \tfrac{1}{2}(x-x_1)\cdot M(x-x_1)\right| \le \tfrac{1}{6}\kappa(|x-x_1|^3 + |t-t_1|^{3/2}).$$

Notice that, since $u_t + \bar{F}(D^2u) = f(x,t)$ in U_T and u is touched from above and below at (x_1,t_1) by polynomials with Hessians equal to M, this implies that $b + \bar{F}(M) = f(x_1,t_1)$. Therefore, defining

$$\phi(x,t) := u(x_1,t_1) + b(t-t_1) + (\xi - p)\cdot(x-x_1) + \tfrac{1}{2}(x-x_1)\cdot M(x-x_1) - \tfrac{1}{6}\kappa(|x-x_1|^3 + |t-t_1|^{3/2}),$$

we have

$$\begin{aligned} u(x_1,t_1) - v(y_1,s_1) - \frac{1}{2\delta}[|x_1-y_1|^2 + (t_1-s_1)^2] \\ \ge \sup_{U_T\times U_T, t\le t_1, s\le s_1} \left\{\phi(x,t) - v(y,s) - \frac{1}{2\delta}[|x-y|^2 + (t-s)^2] - q\cdot(y-y_1)\right\}. \end{aligned} \tag{6-19}$$

To control the right-hand side from below, we consider that, for any $(y,s) \in U_T$ with $s \le s_1$, letting $x = x_1 + y - y_1$ and $t = t_1 + s - s_1 \le t_1$,

$$\begin{aligned} &\sup_{(x,t)\in U_T, t\le t_1} \left\{\phi(x,t) - \frac{1}{2\delta}[|x-y|^2 + (t-s)^2]\right\} \\ &\quad\ge \phi(x_1+y-y_1, t_1+s-s_1) - 12[|x_1-y_1|^2 + (t_1-s_1)^2] \\ &\quad= u(x_1,t_1) + b(s-s_1) + (\xi-p)\cdot(y-y_1) + \tfrac{1}{2}(y-y_1)\cdot M(y-y_1) \\ &\qquad\qquad - \tfrac{1}{6}\kappa(|y-y_1|^3 + |s-s_1|^{3/2}) - \frac{1}{2\delta}[|x_1-y_1|^2 + (t_1-s_1)^2]. \end{aligned} \tag{6-20}$$

Combining (6-19) and (6-20) yields that

$$\begin{aligned} &u(x_1,t_1) - v(y_1,s_1) - \frac{1}{2\delta}[|x_1-y_1|^2 + (t_1-s_1)^2] \\ &\ge \sup_{(y,s)\in U_T, s\le s_1} \Big\{u(x_1,t_1) + b(s-s_1) + (\xi-p)\cdot(y-y_1) + \tfrac{1}{2}(y-y_1)\cdot M(y-y_1) - \tfrac{1}{6}\kappa(|y-y_1|^3 + |s-s_1|^{3/2}) \\ &\qquad\qquad - \frac{1}{2\delta}[|x_1-y_1|^2 + (t_1-s_1)^2] - v(y,s) - q\cdot(y-y_1)\Big\}. \end{aligned}$$

This implies that

$$\begin{aligned} v(y_1,s_1) \le \inf_{(y,s)\in U_T, s\le s_1} \Big\{&v(y,s) - b(s-s_1) - (\xi-p-q)\cdot(y-y_1) \\ &- \tfrac{1}{2}(y-y_1)\cdot M(y-y_1) + \tfrac{1}{6}\kappa(|y-y_1|^3 + |s-s_1|^{3/2})\Big\}. \end{aligned} \tag{6-21}$$

Since $l \le \eta$, choose $M^* \in \mathbb{S}^d$ so that $M \le M^* \le M + C\eta^\sigma\,\mathrm{Id}$ and $l^{-1}M^*$ has integer entries. Using that $\bar{F}$ is uniformly elliptic, $\bar{F}(M^*) \le \bar{F}(M) = f(x_1,t_1) - b$. Let

$$\begin{aligned} \Theta(y,s) := v(y,s) &- b(s-s_1) - (\xi-p-q)\cdot(y-y_1) \\ &- \tfrac{1}{2}(y-y_1)\cdot(M - C\eta^\sigma\,\mathrm{Id})(y-y_1) + \tfrac{1}{6}\kappa(|y-y_1|^3 + |s-s_1|^{3/2}). \end{aligned}$$

By (6-13),

$$
\begin{aligned}
\Theta_s &+ F(M^* + D^2\Theta, y, s) \\
&= v_s - b + \tfrac{1}{4}\kappa|s - s_1|^{1/2} \\
&\qquad + F\Bigl(M^* + D^2 v - M + C\eta^\sigma \operatorname{Id} + \tfrac{1}{2}\kappa|y - y_1| \operatorname{Id} + \tfrac{1}{2}\kappa \frac{(y - y_1) \otimes (y - y_1)}{|y - y_1|}, y, s\Bigr) \\
&\geq v_s - b + F(D^2 v, y, s) - C\bigl(M^* - M + C\eta^\sigma \operatorname{Id} + C\tfrac{1}{2}\kappa|y - y_1| \operatorname{Id}\bigr) \\
&\geq f(y, s) + cA - b - C\eta^\sigma - C\tfrac{1}{2}\kappa|y - y_1| \\
&\geq f(y, s) + cA - b - C\eta^\sigma - C\tfrac{1}{2}(\kappa + 1)|y - y_1| \\
&\geq \bar{F}(M) - CA^6 + cA - C\eta^\sigma - C\tfrac{1}{2}(\kappa + 1)|y - y_1|,
\end{aligned}
$$

where the last line holds by (6-17), using that $\bar{F}(M) = f(x_1, t_1) - b$.

This implies that, in $Q_{cA(\kappa+1)^{-1}}(y_1, s_1)$,

$$\Theta_s + F(M^* + D^2\Theta, y, s) \geq \bar{F}(M) - CA^6 + cA - C\eta^\sigma.$$

In addition, comparing (6-21) and the definition of Θ,

$$\Theta(y_1, s_1) \leq \inf_{(y,s)\in U_T, s\leq s_1} (\Theta - C\eta^\sigma |y - y_1|^2). \tag{6-22}$$

Let (y^*, s^*) be such that $(\eta^{-1}y^*, \eta^{-2}s^*) \in \mathbb{Z}^{d+1}$ and $d[(y^*, s^*), (y_1, s_1)] \leq \sqrt{d}\eta$.

Let

$$G^* := (y^*, s^*) - \eta G_0.$$

Since $(y_1, s_1) \in U_T(\rho)$, we have $d[(y^*, s^*), \partial_p U_T] \geq \rho - \sqrt{d}\eta \geq \sqrt{d}\eta$ so long as $\rho := CA^{1/\alpha} \geq C\eta$ (which is satisfied if $\sigma \leq \alpha$). This implies that $G^* \subseteq U_T$.

We next claim that $G^* \subseteq Q_{cA(\kappa+1)^{-1}}(y_1, s_1)$ for an appropriate choice of κ. Let $\kappa := \eta^{\sigma-1}$ with $\sigma := \bigl((1+4(d+2))/\alpha^2\bigr)^{-1} \leq \alpha$. Since $A \geq C\eta^\sigma$, we may choose the constants so that $cA(\kappa+1)^{-1} \geq \sqrt{d}\eta$. This yields that $G^* \subseteq Q_{cA(\kappa+1)^{-1}}(y_1, s_1)$, as asserted.

Therefore,

$$\Theta_s + F(M^* + D^2\Theta, y, s) \geq \bar{F}(M^*) \quad \text{in } G^*. \tag{6-23}$$

By (6-22), we conclude that

$$\inf_{G^*} \Theta \leq \inf_{\partial_p G^*} \Theta - C\eta^\sigma. \tag{6-24}$$

This implies, by Lemma 2.1 and (6-24), that

$$\mu(G^*, \bar{F}(M^*), M^*) \geq cA^{d+1}$$

and this completes the proof. □

Finally, we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. We prove a rate in probability for the decay of $u - u^\varepsilon$. Fix M_0 and U_T such that $U_T \subset Q_1$ and

$$1 + K_0 + \|g\|_{C^{0,1}(\partial_p U_T)} \leq M_0.$$

We will show that there exists $\beta > 0$ and a random variable $\mathscr{X} : \Omega \to \mathbb{R}$ such that

$$\sup_{U_T}\{u(x,t) - u^\varepsilon(x,t,\omega)\} \leq C[1 + \varepsilon^p \mathscr{X}(\omega)]\varepsilon^\beta.$$

We mention that a rate on $u^\varepsilon - u$ follows by a completely analogous argument for μ^*, so we choose to omit it.

Fix $\varepsilon \in (0, 1)$ and $p < d+2$, and let σ be as in Proposition 6.3. Let α be the α associated with p as in Corollary 5.5 and let $q := \frac{1}{4}p$. Choose m such that

$$\max\{3^{-m/4}, 3^{-m\alpha/(d+1)}\} \leq \varepsilon. \tag{6-25}$$

In the language of Proposition 6.3, let $\eta := 3^{-m\alpha/2(d+1)}$ and choose $l := 3^{-m\alpha/2d}$. Notice that we have that $l \leq \eta \leq \varepsilon^{1/2}$. This implies that, for any $A \geq C\eta^\sigma$,

$$\begin{aligned}\Big\{\omega : \sup_{(x,t)\in U_T} u(x,t) - u^\varepsilon(x,t,\omega) \geq A\Big\}& \\ &\subseteq \bigcup_{(y,s,M)\in \mathscr{I}(A)} \big\{\omega : \mu((y/\varepsilon, s/\varepsilon^2) + \eta\varepsilon^{-1}G_0, \omega, \bar{F}(M), M) \geq cA^{d+1}\big\} \\ &= \bigcup_{(y,s,M)\in \mathscr{I}(A)} \big\{\omega : \mu((y/\varepsilon, s/\varepsilon^2) + G_m, \omega, \bar{F}(M), M) \geq cA^{d+1}\big\},\end{aligned}$$

where

$$\mathscr{I}(A) := \big\{(y, s, M) : (y, s) \in Q_1, (\eta^{-1}y, \eta^{-2}s) \in \mathbb{Z}^{d+1}, |M| \leq 3^{m\alpha/2(d+1)}\big\}.$$

This is possible since $\eta < 1$ and Proposition 6.3 yields that $\sigma < 1$, which implies that $|M| \leq \eta^{\sigma-1} \leq \eta^{-1} \leq 3^{m\alpha/2(d+1)}$. We mention also that $l^{-1}M \in \mathbb{Z}^{d^2} \cap \mathbb{S}^d$.

This implies that

$$\sup_{(x,t)\in U_T}\{u(x,t) - u^\varepsilon(x,t,\omega)\} \leq cA^{d+1} + \mathscr{Y}_m(\omega), \tag{6-26}$$

where

$$\mathscr{Y}_m(\omega) := \big\{\sup \mu((z, r) + G_m, \omega, \bar{F}(M), M) : (z\varepsilon^{-1}, r\varepsilon^{-2}, M) \in \mathscr{I}(A)\big\}. \tag{6-27}$$

To find the number of elements in $\mathscr{I}(A)$, consider that, since $\eta^{-1}z \in \mathbb{Z}^d \cap Q_{1/\varepsilon}$ and $\eta^{-2}s \in \mathbb{Z} \cap [0, 1/\varepsilon^2]$, there are $(\varepsilon\eta)^{-(d+2)}$ choices for (z, s). This implies that there are at most $3^{3m\alpha}$ choices. For the matrices, consider that, since $3^{m\alpha/2d}M \in \mathbb{Z}^{d^2} \cap \mathbb{S}^d$ and $|M| \leq 3^{m\alpha/2(d+1)}$, this implies that there are at most $3^{m\alpha(d+1)}$ terms. In total, there are $3^{m\alpha(d+4)}$ combinations to choose from in $\mathscr{I}(A)$.

By Corollary 5.5, for each $(z, r, M) \in \mathscr{I}(A)$,

$$\mathbb{P}\big[(z, r) + \mu(G_m, \omega, \bar{F}(M), M) \geq (1 + |M|)^{d+1}3^{-m\alpha}\tau\big] \leq C\exp(-c3^{mp}\tau).$$

Since $|M|^{d+1} \leq 3^{m\alpha/2}$, this implies that

$$\mathbb{P}\big[(z,r)+\mu(G_m,\omega,\overline{F}(M),M) \geq 3^{-m\alpha/2}\tau\big] \leq \exp(-c3^{mp}\tau).$$

Using a union bound and summing over all of the terms in $\mathcal{I}(A)$,

$$\mathbb{P}\big[\mathcal{Y}_m(\omega) \geq 3^{-m\alpha/2}\tau\big] \leq C3^{m\alpha(d+4)}\exp(-c3^{mp}\tau) \leq C\exp(-c3^{mp}\tau).$$

Replacing τ by $\tau+1$, we have that, for all $\tau \geq 0$,

$$\mathbb{P}\big[(3^{m\alpha/2}\mathcal{Y}_m(\omega)-1)_+ \geq \tau\big] \leq C\exp(-c3^{mp}\tau).$$

Replacing again $\tau \to 3^{-mq}\tau$ yields that

$$\mathbb{P}\big[3^{mq}3^{m\alpha/2}(\mathcal{Y}_m(\omega)-1)_+ \geq \tau\big] \leq C\exp(-c3^{m(p-q)}\tau).$$

Summing over m and using that $p > q$, this implies that, for all $\tau \geq 0$,

$$\mathbb{P}\big[\sup_m\{3^{mq}3^{m\alpha/2}(\mathcal{Y}_m(\omega)-1)_+\} \geq \tau\big] \leq \sum_m \mathbb{P}\big[3^{mq}3^{m\alpha/2}(\mathcal{Y}_m(\omega)-1)_+ \geq \tau\big] \leq C\exp(-c\tau). \tag{6-28}$$

Letting

$$\mathcal{X}(\omega) := \sup_m\{3^{mq}(3^{m\alpha/2}\mathcal{Y}_m(\omega)-1)_+\} \tag{6-29}$$

and integrating (6-28) in τ yields that

$$\mathbb{E}\big[\exp(\mathcal{X}(\omega))\big] \leq C. \tag{6-30}$$

This implies that

$$\sup_{(x,t)\in U_T}\{u(x,t)-u^\varepsilon(x,t,\omega)\} \leq C\eta^{\sigma(d+1)} + C(3^{-mq}\mathcal{X}(\omega)+1)3^{-m\alpha/2} \leq C(1+\varepsilon^p\mathcal{X}(\omega))\varepsilon^\beta$$

for some choice of β, where $\beta(\lambda,\Lambda,d,p)$. □

Acknowledgements

Part of this article appeared in Lin's doctoral thesis. Both authors would like to thank Scott Armstrong and Takis Souganidis for useful discussions. Lin was partially supported by NSF grants DGE-1144082 and DMS-1147523. Smart was partially supported by NSF grant DMS-1461988 and the Sloan Foundation. This collaboration took place at the Mittag-Leffler Institute.

References

[Akcoglu and Krengel 1980] M. A. Akcoglu and U. Krengel, "Ergodic theorems for superadditive processes", *C. R. Math. Rep. Acad. Sci. Canada* **2**:4 (1980), 175–179. MR 81i:60056 Zbl 0439.60025

[Armstong and Daniel 2015] S. N. Armstong and J. P. Daniel, personal communication, 2015.

[Armstrong and Mourrat 2015] S. N. Armstrong and J. C. Mourrat, "Lipschitz regularity for elliptic equations with random coefficients", *Arch. Ration. Mech. Anal.* (online publication July 2015).

[Armstrong and Smart 2014a] S. N. Armstrong and C. K. Smart, "Quantitative stochastic homogenization of convex integral functionals", preprint, 2014. To appear in *Ann. Sci. Éc. Norm. Supér.* arXiv 1406.0996v3

[Armstrong and Smart 2014b] S. N. Armstrong and C. K. Smart, "Quantitative stochastic homogenization of elliptic equations in nondivergence form", *Arch. Ration. Mech. Anal.* **214**:3 (2014), 867–911. MR 3269637 Zbl 1304.35714

[Caffarelli and Souganidis 2010] L. A. Caffarelli and P. E. Souganidis, "Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media", *Invent. Math.* **180**:2 (2010), 301–360. MR 2011c:35041 Zbl 1192.35048

[Caffarelli et al. 2005] L. A. Caffarelli, P. E. Souganidis, and L. Wang, "Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media", *Comm. Pure Appl. Math.* **58**:3 (2005), 319–361. MR 2006b:35016 Zbl 1063.35025

[Crandall 1997] M. G. Crandall, "Viscosity solutions: a primer", pp. 1–43 in *Viscosity solutions and applications* (Montecatini Terme, 1995), edited by I. Capuzzo Dolcetta and P. L. Lions, Lecture Notes in Math. **1660**, Springer, Berlin, 1997. MR 98g:35034 Zbl 0901.49026

[Crandall et al. 1992] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.* (*N.S.*) **27**:1 (1992), 1–67. MR 92j:35050 Zbl 0755.35015

[Crandall et al. 1999] M. G. Crandall, M. Kocan, P. L. Lions, and A. Święch, "Existence results for boundary problems for uniformly elliptic and parabolic fully nonlinear equations", *Electron. J. Differential Equations* **1999**:24 (1999), 1–20. MR 2000f:35052 Zbl 0927.35029

[Daniel 2015] J. P. Daniel, "Quadratic expansions and partial regularity for fully nonlinear uniformly parabolic equations", *Calc. Var. Partial Differential Equations* **54**:1 (2015), 183–216. MR 3385158

[Daskalopoulos and Savin 2012] P. Daskalopoulos and O. Savin, "$C^{1,\alpha}$ regularity of solutions to parabolic Monge–Ampére equations", *Amer. J. Math.* **134**:4 (2012), 1051–1087. MR 2956257 Zbl 1258.35133

[Evans 1992] L. C. Evans, "Periodic homogenisation of certain fully nonlinear partial differential equations", *Proc. Roy. Soc. Edinburgh Sect. A* **120**:3–4 (1992), 245–265. MR 93a:35016 Zbl 0796.35011

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Fischer and Otto 2015] J. Fischer and F. Otto, "A higher-order large-scale regularity theory for random elliptic operators", preprint, 2015. arXiv 1503.07578

[Gloria et al. 2014] A. Gloria, S. Neukamm, and F. Otto, "A regularity theory for random elliptic operators", preprint, 2014. arXiv 1409.2678v2

[Gutiérrez and Huang 1998] C. E. Gutiérrez and Q. Huang, "A generalization of a theorem by Calabi to the parabolic Monge–Ampère equation", *Indiana Univ. Math. J.* **47**:4 (1998), 1459–1480. MR 2000a:35105 Zbl 0926.35053

[Gutiérrez and Huang 2001] C. E. Gutiérrez and Q. Huang, "$W^{2,p}$ estimates for the parabolic Monge–Ampère equation", *Arch. Ration. Mech. Anal.* **159**:2 (2001), 137–177. MR 2002h:35118 Zbl 0992.35020

[Imbert and Silvestre 2012] C. Imbert and L. Silvestre, "Lecture notes on fully nonlinear parabolic equations", lecture notes, 2012, available at http://blog-cyrilimbert.net/2012/06/20/lecture-notes-on-fully-nonlinear-parabolic-equations/.

[Krylov 1976] N. V. Krylov, "Sequences of convex functions, and estimates of the maximum of the solution of a parabolic equation", *Sibirsk. Mat. Ž.* **17**:2 (1976), 290–303. In Russian; translated in *Sib. Math. J.* **17**:2 (1976), 226–236. MR 54 #8033 Zbl 0362.35038

[Krylov 1982] N. V. Krylov, "Boundedly inhomogeneous elliptic and parabolic equations", *Izv. Akad. Nauk SSSR Ser. Mat.* **46**:3 (1982), 487–523, 670. In Russian; translated in *Math. USSR Izv.* **20**:3 (1983), 459–492. MR 84a:35091 Zbl 03806019

[Lin 2015] J. Lin, "On the stochastic homogenization for fully nonlinear uniformly parabolic equations in stationary ergodic spatio-temporal media", *J. Differential Equations* **258**:3 (2015), 796–845. Zbl 06371761

[Papanicolaou and Varadhan 1982] G. C. Papanicolaou and S. R. S. Varadhan, "Diffusions with random coefficients", pp. 547–552 in *Statistics and probability: essays in honor of C. R. Rao*, edited by G. Kallianpur et al., North-Holland, Amsterdam, 1982. MR 85e:60082 Zbl 0486.60076

[Tso 1985] K. Tso, "On an Aleksandrov–Bakel′man type maximum principle for second-order parabolic equations", *Comm. Partial Differential Equations* **10**:5 (1985), 543–553. MR 87f:35031 Zbl 0581.35027

[Turanova 2015] O. Turanova, "Error estimates for approximations of nonlinear uniformly parabolic equations", *Nonlinear Differential Equations Appl.* **22**:3 (2015), 345–389. MR 3349798 Zbl 06449239

[Wang 1992] L. Wang, "On the regularity theory of fully nonlinear parabolic equations, I", *Comm. Pure Appl. Math.* **45**:1 (1992), 27–76. MR 92m:35126 Zbl 0832.35025

[Yurinskiĭ 1982] V. V. Yurinskiĭ, "Averaging of nondivergence random elliptic operators", pp. 126–138, 207 in *Limit theorems of probability theory and related questions*, edited by S. L. Sobolev, Trudy Inst. Mat. **1**, "Nauka" Sibirsk. Otdel., Novosibirsk, 1982. MR 84d:60097 Zbl 0518.35088

[Yurinskiĭ 1988] V. V. Yurinskiĭ, "On the error of averaging of multidimensional diffusions", *Teor. Veroyatnost. i Primenen.* **33**:1 (1988), 14–24. In Russian; translated in *Theory Probab. Appl.* **33**:1 (1988) 1–21. MR 89e:60152 Zbl 0672.60074

[Yurinskiĭ 1991] V. Yurinskiĭ, "Homogenization error estimates for random elliptic operators", pp. 285–291 in *Mathematics of random media* (Blacksburg, VA, 1989), edited by W. E. Kohler and B. S. White, Lectures in Appl. Math. **27**, Amer. Math. Soc., Providence, RI, 1991. MR 92j:60079 Zbl 0729.60060

Received 22 Jan 2015. Revised 7 May 2015. Accepted 24 Jun 2015.

JESSICA LIN: jessica@math.wisc.edu
Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53706, United States

CHARLES K. SMART: smart@math.cornell.edu
Department of Mathematics, Cornell University, 401 Malott Hall, Ithaca, NY 14853, 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 8 No. 6 2015
