Analysis & PDE

Volume 9 No. 2 2016

msp

Analysis & PDE
msp.org/apde

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

The subscription price for 2016 is US $235/year for the electronic version, and $430/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. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.259

RESONANCES FOR LARGE ONE-DIMENSIONAL “ERGODIC” SYSTEMS

FRÉDÉRIC KLOPP

Dedicated to Johannes Sjöstrand on the occasion of his seventieth birthday.

The present paper is devoted to the study of resonances for one-dimensional quantum systems with a potential that is the restriction to some large box of an ergodic potential. For discrete models, both on a half-line and on the whole line, we study the distributions of the resonances in the limit when the size of the box goes to infinity. For periodic and random potentials, we analyze how the spectral theory of the limit operator influences the distribution of the resonances.

Dans cet article, nous étudions les résonances d’un système unidimensionnel plongé dans un potentiel qui est la restriction à un grand intervalle d’un potentiel ergodique. Pour des modèles discrets sur la droite et la demie droite, nous étudions la distribution des résonances dans la limite de la taille de boîte infinie. Pour des potentiels périodiques et aléatoires, nous analysons l’influence de la théorie spectrale de l’opérateur limite sur la distribution des résonances.

0. Introduction

Consider $V:\mathbb{Z}\to\mathbb{R}$ a bounded potential and, on $\ell^2(\mathbb{Z})$, the Schrödinger operator $H=-\Delta+V$ defined by

$$(Hu)(n)=u(n+1)+u(n-1)+V(n)u(n)\quad\text{for all } n\in\mathbb{Z}$$

for $u\in\ell^2(\mathbb{Z})$.

The potentials V we will deal with are of two types:

It is a pleasure to thank N. Filonov for interesting discussions at the early stages of this work and T. T. Phong, C. Shirley and M. Vogel for pointing out misprints in previous versions of the article. This work was partially supported by the grant ANR-08-BLAN-0261-01.
MSC2010: 35B34, 47B80, 47H40, 60H25, 82B44.
Keywords: resonances, random operators, periodic operators.

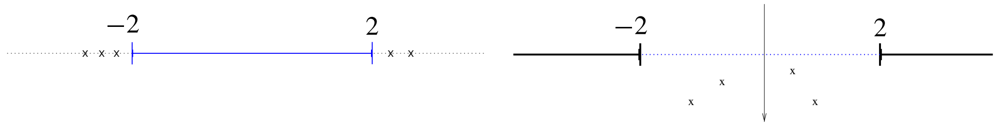


Figure 1. The meromorphic continuation.

- V periodic;
- $V = V_\omega$, the random Anderson model, i.e., the entries of the diagonal matrix V are independent, identically distributed, nonconstant random variables.

The spectral theory of such models has been studied extensively (see, e.g., [Kirsch 2008]) and it is well known that

- when V is periodic, the spectrum of H is purely absolutely continuous;
- when $V = V_\omega$ is random, the spectrum of H is almost surely pure point, i.e., the operator only has eigenvalues; moreover, the eigenfunctions decay exponentially at infinity.

Pick $L \in \mathbb{N}^*$. The main object of our study is the operator

$$H_L = -\Delta + V\mathbf{1}_{[\![-L+1,L]\!]} \tag{0-1}$$

when L is large. Here, $[\![-L+1, L]\!]$ is the integer interval $\{-L+1, \dots, L\}$, and $\mathbf{1}_{[\![a,b]\!]}(n) = 1$ if $a \le n \le b$ and $\mathbf{1}_{[\![a,b]\!]}(n) = 0$ if not.

For L large, the operator H_L is a simple Hamiltonian modeling a large sample of periodic or random material in the void. It is well known in this case (see, e.g., [Zworski 2002]) that not only is the spectrum of H_L of importance but also its (quantum) resonances, which we will now define.

As $V\mathbf{1}_{[\![-L+1,L]\!]}$ has finite rank, the essential spectrum of H_L is the same as that of the discrete Laplace operator, that is, $[-2, 2]$, and it is purely absolutely continuous. Outside this absolutely continuous spectrum, H_L has only discrete eigenvalues associated to exponentially decaying eigenfunctions.

We are interested in the resonances of the operator H_L in the limit when $L \to +\infty$. They are defined to be the poles of the meromorphic continuation of the resolvent of H_L through $(-2, 2)$, the continuous spectrum of H_L (see Figure 1, Theorem 1.3 and, e.g., [loc. cit.]). The resonances widths, that is, their imaginary part, play an important role in the large time behavior of e^{-itH_L}, especially the resonances of smallest width that give the leading order contribution (see [loc. cit.]).

Quantum resonances are basic objects in quantum theory. They have been the focus of a vast number of studies, both mathematical and physical (see, e.g., [loc. cit.] and references therein). Our purpose here is to study the resonances of H_L in the asymptotic regime $L \to +\infty$. As $L \to +\infty$, H_L converges to H in the strong resolvent sense. Thus, it is natural to expect that the differences in the spectral nature between the cases V periodic and V random should reflect into differences in the behavior of the resonances in both cases. We shall see below that this is the case. To illustrate this as simply as possible, we begin by stating three theorems, one for periodic potentials and two for random potentials, that underline these different behaviors. These results can be considered as paradigmatic for our main results, presented in Section 1.

The scattering theory or the closely related questions of resonances for the operator (0-1) or for closely related one-dimensional models have already been discussed in various works, both in the mathematical and physical literature (see, e.g., [Faris and Tsay 1989; 1994; Lifshits et al. 1988; Kunz and Shapiro 2006; Texier and Comtet 1999; Comtet and Texier 1997; Kunz and Shapiro 2008; Barra and Gaspard 1999; Kottos 2005; Titov and Fyodorov 2000]). We will make more comments on the literature as we develop our results in Section 1.

0A. ***When V is periodic.*** Assume that V is p-periodic ($p \in \mathbb{N}^*$) and does not vanish identically. Consider $H = -\Delta + V$ and let $\Sigma_{\mathbb{Z}}$ be its spectrum, $\Sigma_{\mathbb{Z}}^{\circ}$ be its interior and $E \mapsto N(E)$ be its integrated density of states, i.e., the number of states of the system per unit of volume below energy E (see Section 1B and, e.g., [Teschl 2000] for precise definitions and details).

Theorem 0.1. *There exist*

- $\mathfrak{D}$, *a discrete (possibly empty) set of energies in* $(-2, 2) \cap \Sigma_{\mathbb{Z}}^{\circ}$,
- *a function h that is real analytic in a complex neighborhood of* $(-2, 2)$ *and that does vanish on* $(-2, 2) \setminus \mathfrak{D}$

such that, for $I \subset (-2, 2) \setminus \mathfrak{D}$ *a compact interval such that either* $I \cap \Sigma_{\mathbb{Z}} = \varnothing$ *or* $I \subset \Sigma_{\mathbb{Z}}^{\circ}$, *there exists* $c_0 > 0$ *such that, for L sufficiently large with* $L \in p\mathbb{N}$, *one has:*

- *If* $I \cap \Sigma_{\mathbb{Z}} = \varnothing$, *then* H_L *has no resonance in* $I + i[-c_0, 0]$.
- *If* $I \subset \Sigma_{\mathbb{Z}}^{\circ}$, *one has:*
  - *There are plenty of resonances in* $I + i[-c_0, 0]$; *more precisely,*

$$\frac{1}{2L}\#\{z \in I + i[-c_0, 0] \mid z \text{ a resonance of } H_L\} = \int_I dN(E) + o(1), \tag{0-2}$$

  where $o(1) \to 0$ *as* $L \to +\infty$.
  - *Let* $(z_j)_j$ *be the resonances of* H_L *in* $I + i[-c_0, 0]$ *ordered by increasing real part; then*

$$L \cdot \mathrm{Re}(z_{j+1} - z_j) \asymp 1 \quad \text{and} \quad L \cdot \mathrm{Im}\, z_j = h(\mathrm{Re}\, z_j) + o(1), \tag{0-3}$$

  the estimates in (0-3) *being uniform for all the resonances in* $I + i[-c_0, 0]$ *when* $L \to +\infty$.

After rescaling their width by L, resonances are nicely interspaced points lying on an analytic curve (see Figure 2). We give a more precise description of the resonances in Theorem 1.7 and Propositions 1.8 and 1.9. In particular, we describe the set of energies $\mathfrak{D}$ and the resonances near these energies: they lie further away from the real axis, the maximal distance being of order $L^{-1} \log L$ (see Figure 3). Theorem 0.1 only describes the resonances closest to the real axis. In Section 1B, we also give results on the resonances located deeper in the lower half of the complex plane.

0B. ***When V is random.*** Assume now that $V = V_\omega$ is the Anderson potential, i.e., its entries are i.i.d. and distributed uniformly on [0, 1] for concreteness. Consider $H = -\Delta + V_\omega$. Let Σ be its almost sure spectrum (see, e.g., [Pastur and Figotin 1992] for this and the following notions), $E \mapsto n(E)$ its density of states (i.e., the derivative of the integrated density of states; see also Section 1B) and $E \mapsto \rho(E)$

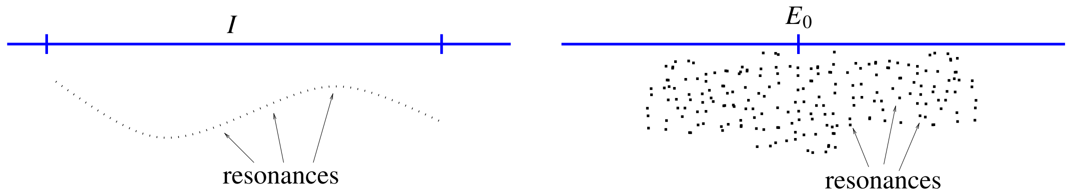


Figure 2. The rescaled resonances for the periodic (left) and the random (right) potential.

its Lyapunov exponent (see also Section 1C). The Lyapunov exponent is known to be continuous and positive; the density of states satisfies $n(E) > 0$ for a.e. $E \in \Sigma$ (see, e.g., [Bougerol and Lacroix 1985]).

Define $H_{\omega,L} := -\Delta + V_\omega \mathbf{1}_{[\![-L+1,L]\!]}$. We prove:

Theorem 0.2. *Pick $I \subset (-2, 2)$ a compact interval.*

- *If $I \cap \Sigma = \varnothing$ then there exists $c_I > 0$ such that ω-a.s., for L sufficiently large,*
$$\{z \text{ a resonance of } H_{\omega,L} \text{ in } I + i(-c_I, 0]\} = \varnothing.$$
- *If $I \subset \Sigma^\circ$ then, for any $c > 0$, ω-a.s. one has*
$$\lim_{L\to+\infty} \frac{1}{L} \#\{z \text{ a resonance of } H_{\omega,L} \text{ in } I + i(-\infty, -e^{-2cL}]\} = \int_I \min\left(\frac{c}{\rho(E)},\ 1\right) n(E)\, dE.$$

As the first statement of Theorem 0.2 is clear, let us discuss the second. Define $c_+ := \max_{E\in I} \rho(E)$. For $c \ge c_+$, ω-a.s. for L large the number of resonances in the strip $\{\operatorname{Re} z \in I,\ \operatorname{Im} z \le -e^{-2cL}\}$ is approximately $2L \int_I n(E)\, dE$; thus, in $\{\operatorname{Re} z \in I,\ -e^{2c_+L} \le \operatorname{Im} z < 0\}$, one finds at most $o(L)$ resonances. We shall see that, for $\delta > 0$, ω-a.s. for L large the strip $\{\operatorname{Re} z \in I,\ -e^{(2c_++\delta)L} \le \operatorname{Im} z < 0\}$ actually contains no resonances (see Theorem 1.13).

Define $c_- := \min_{E\in I} \rho(E)$. For $c \le c_-$, ω-a.s. for L large the strip $\{\operatorname{Re} z \in I,\ \operatorname{Im} z \le -e^{-2cL}\}$ contains approximately $2cL \int_I n(E)/\rho(E)\, dE$ resonances. We shall see that, for $\kappa \in [0, 1)$, the number of resonances in the strip $\{\operatorname{Re} z \in I,\ \operatorname{Im} z \le -e^{-L^\kappa}\}$ is $O(L^\kappa)$, thus $o(L)$ (see Theorem 1.17).

One can also describe the resonances locally. Fix $E_0 \in (-2, 2) \cap \Sigma^\circ$ such that $n(E_0) > 0$. Let $(z_l^L(\omega))_l$ be the resonances of $H_{\omega,L}$. We first rescale them. Define

$$x_l^L(\omega) = 2Ln(E_0)(\operatorname{Re} z_l^L(\omega) - E_0) \quad \text{and} \quad y_l^L(\omega) = -\frac{1}{2L\rho(E_0)} \log|\operatorname{Im} z_l^L(\omega)|. \tag{0-4}$$

Consider now the two-dimensional point process

$$\xi_L(E_0, \omega) = \sum_{z_l^L \text{ resonances of } H_{\omega,L}} \delta_{(x_l^L(\omega),\, y_l^L(\omega))}.$$

We prove:

Theorem 0.3. *The point process ξ_L converges weakly to a Poisson process of intensity* 1 *in* $\mathbb{R} \times [0, 1]$.

In the random case, the structure of the (properly rescaled) resonances is quite different from that in the periodic case (see Figure 2). The real parts of the resonances are scaled in such a way that their average spacing becomes of order one. By Theorem 0.2, the imaginary parts are typically exponentially small (in L); when the resonances are rescaled as in (0-4), their imaginary parts are rewritten on a logarithmic scale so as to become of order 1 too. Once rescaled in this way, the local picture of the resonances of $H_{\omega,L}$ is that of a two-dimensional cloud of Poisson points (see the right-hand side of Figure 2).

Theorem 0.3 is the analogue for resonances of the well-known result on the distribution of eigenvalues and localization centers for the Anderson model in the localized phase (see, e.g., [Minami 1996; Killip and Nakano 2007; Germinet and Klopp 2014]).

As in the case of the periodic potential, Theorem 0.3 only describes the resonances closest to the real axis. In Section 1C, we also give results on resonances located deeper in the lower half of the complex plane. Up to distances of order $L^{-\infty}$ to the real axis, the cloud of resonances (once properly rescaled) will have the same Poissonian behavior as described above (see Theorem 1.10).

Besides proving Theorems 0.1 and 0.3, the goal of the paper is to describe the statistical properties of the resonances and relate them (the distribution of the resonances and of the widths) to the spectral characteristics of $H = -\Delta + V$, and possibly to the distribution of its eigenvalues (see, e.g., [Germinet and Klopp 2011]).

As they can be analyzed in a very similar way, we will discuss three models:

- The model H_L defined above.
- Its analogue on the half-line $\mathbb{N}$, i.e., on H_L, we impose an additional Dirichlet boundary condition at 0.
- The “half-infinite” model on $\ell^2(\mathbb{Z})$, that is,
$$H^\infty = -\Delta + W, \quad \text{where} \quad \begin{cases} W(n) = 0 & \text{for } n \geq 0, \\ W(n) = V(n) & \text{for } n \leq -1, \end{cases} \tag{0-5}$$
where V is chosen as above, periodic or random.

Though in the present paper we restrict ourselves to discrete models, it is clear that continuous one-dimensional models can be dealt with essentially using the methods developed here.

1. The main results

We now turn to our main results, a number of which were announced in [Klopp 2012]. Pick $V : \mathbb{Z} \to \mathbb{R}$ a bounded potential and, for $L \in \mathbb{N}$, consider the operators

- $H_L^{\mathbb{Z}} = -\Delta + V\mathbf{1}_{[\![0,L]\!]}$ on $\ell^2(\mathbb{Z})$;
- $H_L^{\mathbb{N}} = -\Delta + V\mathbf{1}_{[\![0,L]\!]}$ on $\ell^2(\mathbb{N})$ with Dirichlet boundary conditions at 0;
- H^∞, defined in (0-5).

Remark 1.1. Here, by “Dirichlet boundary condition at 0”, we mean that $H_L^{\mathbb{N}}$ is the operator $H_L^{\mathbb{Z}}$ restricted to the subspace $\ell^2(\mathbb{N})$, i.e., if $\Pi : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{N})$ is the orthogonal projector on $\ell^2(\mathbb{N})$, one has $H_L^{\mathbb{N}} = \Pi H_L^{\mathbb{Z}} \Pi$. In the literature, this is sometime called “Dirichlet boundary condition at -1” (see, e.g., [Teschl 2000]).

For the sake of simplicity, in the half-line case we only consider Dirichlet boundary conditions at 0. But the proofs show that these are not crucial; any selfadjoint boundary condition at 0 would do and, mutatis mutandis, the results would be the same.

Note also that by a shift of the potential V, replacing L by $L+L'$, studying $H_L^{\mathbb{Z}}$ is equivalent to studying $H_{L,L'} = -\Delta + V\mathbf{1}_{[\![-L',L]\!]}$ on $\ell^2(\mathbb{Z})$. Thus, to derive the results of Section 0 from those in the present section, it suffices to consider the models above, in particular $H_L^{\mathbb{Z}}$.

For the models $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$, we start with a discussion of the existence of a meromorphic continuation of the resolvent, then study the resonances when V is periodic and finally turn to the case when V is random.

As H^∞ is not a relatively compact perturbation of the Laplacian, the existence of a meromorphic continuation of its resolvent depends on the nature of V; so, it will be discussed when specializing to V periodic or random.

Remark 1.2 (notations). In the sequel, we write $a \lesssim b$ if for some $C > 0$ (independent of the parameters coming into a or b) one has $a \le Cb$. We write $a \asymp b$ if $a \lesssim b$ and $b \lesssim a$.

1A. *The meromorphic continuation of the resolvent.* One proves the well-known and simple:

Theorem 1.3. *The operator-valued functions $z \mapsto (z - H_L^{\mathbb{N}})^{-1}$ and $z \mapsto (z - H_L^{\mathbb{Z}})^{-1}$ for $z \in \mathbb{C}^+$ admit a meromorphic continuation from $\mathbb{C}^+$ to $\mathbb{C} \setminus ((-\infty, -2] \cup [2, +\infty))$ through $(-2, 2)$* (see Figure 1) *with values in the operators from l^2_{comp} to l^2_{loc}.*

Moreover, the number of poles of each of these meromorphic continuations in the lower half-plane is at most equal to L.

The resonances are defined to be the poles of this meromorphic continuation (see Figure 1).

1B. *The periodic case.* We assume that, for some $p > 0$, one has

$$V_{n+p} = V_n \quad \text{for all } n \ge 0. \tag{1-1}$$

Let $\Sigma_{\mathbb{N}}$ be the spectrum of $H^{\mathbb{N}} = -\Delta + V$ acting on $\ell^2(\mathbb{N})$ with Dirichlet boundary condition at 0 and $\Sigma_{\mathbb{Z}}$ be the spectrum of $H^{\mathbb{Z}} = -\Delta + V$ acting on $\ell^2(\mathbb{Z})$. One has the following description for these spectra:

- $\Sigma_{\mathbb{Z}}$ is a union of intervals, i.e., $\Sigma_{\mathbb{Z}} := \sigma(H) = \bigcup_{j=1}^p [E_j^-, E_j^+]$, where $E_j^- < E_j^+$ $(1 \le j \le p)$ and $a_{j-1}^+ \le E_j^-$ $(2 \le j \le p)$ (see, e.g., [van Moerbeke 1976]); the spectrum of $H^{\mathbb{Z}}$ is purely absolutely continuous and the spectral resolution can be obtained via a Bloch–Floquet decomposition (see, e.g., [loc. cit.]).
- On $\ell^2(\mathbb{N})$ (see, e.g., [Pavlov 1994]), one has
  - $\Sigma_{\mathbb{N}} = \Sigma_{\mathbb{Z}} \cup \{v_j \mid 1 \le j \le n\}$ and $\Sigma_{\mathbb{Z}}$ is the absolutely continuous spectrum of H;
  - the $(v_j)_{0 \le j \le n}$ are isolated simple eigenvalues associated to exponentially decaying eigenfunctions.

It may happen that some of the gaps are closed, i.e., that the number of connected components of $\Sigma_{\mathbb{Z}}$ be strictly less than p. There still is a natural way to write $\Sigma_{\mathbb{Z}} := \sigma(H) = \bigcup_{j=1}^{p}[E_j^-, E_j^+]$ (see Section 4A1), but in this case, for some of the j, one has $E_{j-1}^+ = E_j^-$; we shall call the energies $E_{j-1}^+ = E_j^-$ *closed gaps* (see Definition 4.5). The existence of closed gaps is nongeneric (see [van Moerbeke 1976]).

The operators $H^\bullet$ (for $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$) admit an integrated density of states defined by

$$N(E) = \lim_{L\to+\infty} \frac{\#\{\text{eigenvalues of } (-\Delta+V)|_{[\![-L,L]\!]\cap\bullet} \text{ in } (-\infty, E]\}}{\#([\![-L, L]\!]\cap\bullet)}. \tag{1-2}$$

Here, the restriction of $-\Delta+V$ to $[\![-L, L]\!]\cap\bullet$ is taken with Dirichlet boundary conditions; this is for concreteness as it is known that, in the limit $L\to+\infty$, other selfadjoint boundary conditions would yield the same result for the limit (1-2).

The integrated density of states is the same for $H^{\mathbb{N}}$ and $H^{\mathbb{Z}}$ (see, e.g., [Pastur and Figotin 1992]). It defines the distribution function of some probability measure on $\Sigma_{\mathbb{Z}}$ that is real analytic on $\Sigma_{\mathbb{Z}}^\circ$. Let n denote the density of states of $H^{\mathbb{N}}$ and $H^{\mathbb{Z}}$, that is, $n(E) = dN(E)/dE$.

Remark 1.4. When L gets large, as $H_L^{\mathbb{N}}$ tends to $H^{\mathbb{N}}$ in the strong resolvent sense, interesting phenomena for the resonances of $H_L^{\mathbb{N}}$ should take place near energies in $\Sigma_{\mathbb{N}}$.

Define τ_k to be the shift by k steps to the left, that is, $\tau_k V(\cdot) = V(\cdot + k)$. Then, for $(\ell_L)_L$ such that $l_L \to +\infty$ and $L - \ell_L \to +\infty$ when $L \to +\infty$, $\tau_{l_L}^* H_L^{\mathbb{Z}} \tau_{l_L}$ tend to $H^{\mathbb{Z}}$ in the strong resolvent sense. Thus, interesting phenomena for the resonances of $H_L^{\mathbb{Z}}$ should take place near energies in $\Sigma_{\mathbb{Z}}$.

1B1. *Resonance-free regions.* We start with a description of resonance-free regions near the real axis. To this end, we introduce some operators on the positive and the negative half-lattice.

Above we have defined $H_{\mathbb{N}}$; we shall need another auxiliary operator. On $\ell^2(\mathbb{Z}_-)$ (where $\mathbb{Z}_- = \{n \le 0\}$), consider the operator $H_k^- = -\Delta + \tau_k V$ with Dirichlet boundary condition at 0 (where τ_k is defined to be the shift by k steps to the left, that is, $\tau_k V(\cdot) = V(\cdot + k)$). Let $\Sigma_k^- = \sigma(H_k^-)$.

As is the case for $H^{\mathbb{N}}$, one knows that $\sigma_{\text{ess}}(H_k^-) = \Sigma_{\mathbb{Z}}$ and that $\sigma_{\text{ess}}(H_k^-)$ is purely absolutely continuous (see, e.g., [Teschl 2000, Chapter 7]). H_k^- may also have discrete eigenvalues in $\mathbb{R} \setminus \Sigma_{\mathbb{Z}}$.

We prove:

Theorem 1.5. *Let I be a compact interval in* $(-2, 2)$.

(1) *If $I \subset \mathbb{R}\setminus\Sigma_{\mathbb{N}}$ (resp. $I \subset \mathbb{R}\setminus\Sigma_{\mathbb{Z}}$), then there exists $c > 0$ such that, for L sufficiently large, $H_L^{\mathbb{N}}$ (resp. $H_L^{\mathbb{Z}}$) has no resonances in the rectangle* $\{\operatorname{Re} z \in I,\ \operatorname{Im} z \in [-c, 0]\}$.

(2) *If $I \subset \Sigma_{\mathbb{Z}}$, then there exists $c > 0$ such that, for L sufficiently large, $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$ have no resonances in the rectangle* $\{\operatorname{Re} z \in I,\ \operatorname{Im} z \in [-c/L, 0]\}$.

(3) *Fix $0 \le k \le p-1$ and assume the compact interval I is such that $\{v_j\} = I^\circ \cap \Sigma_{\mathbb{N}} = I \cap \Sigma_{\mathbb{N}}$ and $I \cap \Sigma_{\mathbb{Z}} = \varnothing$* (the $(v_j)_j$ are as defined in the beginning of Section 1B).

 (a) *If $I \cap \Sigma_k^- = \varnothing$ then there exists $c > 0$ such that, for L sufficiently large with $L \equiv k \bmod p$, $H_L^{\mathbb{N}}$ has a unique resonance in the rectangle $\{\operatorname{Re} z \in I,\ -c \le \operatorname{Im} z \le 0\}$; moreover, this resonance, say z_j, is simple and satisfies $\operatorname{Im} z_j \asymp -e^{-\rho_j L}$ and $|z_j - \lambda_j| \asymp e^{-\rho_j L}$ for some $\rho_j > 0$ independent of L.*

(b) *If* $I \cap \Sigma_k^- \neq \varnothing$ *then there exists* $c > 0$ *such that, for* L *sufficiently large with* $L \equiv k \mod p$, $H_L^{\mathbb{N}}$ *has no resonance in the rectangle* $\{\operatorname{Re} z \in I,\ -c \le \operatorname{Im} z \le 0\}$.

So, below the spectral interval $(-2, 2)$, there exists a resonance-free region of width at least of order L^{-1}. For $H_L^{\mathbb{N}}$, if $L \equiv k \mod p$ each discrete eigenvalue of $H^{\mathbb{N}}$ that is not an eigenvalue of H_k^- generates a resonance for $H_L^{\mathbb{N}}$ exponentially close to the real axis (when L is large). When the eigenvalue of H_k^- is also an eigenvalue of $H^{\mathbb{N}} = H_0^+$; it may also generate a resonance but only much further away in the complex plane, at least at a distance of order 1 to the real axis.

In case (3a) of Theorem 1.5, one can give an asymptotic expansion for the resonances (see Section 5B1).

We now turn to the description of the resonances of $H_L^\bullet$ near $[-2, 2]$. To this end, it will be useful to introduce a number of auxiliary functions and operators.

1B2. *Some auxiliary functions.* To H_k^- defined above, we associate N_k^-, the distribution function of its spectral measure (which is a probability measure), i.e., for $\varphi \in \mathscr{C}_0^\infty(\mathbb{R})$, we define $\int_{\mathbb{R}} \varphi(\lambda)\, dN_k^-(\lambda) := \varphi(H_k^-)(0, 0)$, where $(\varphi(H_k^-)(x, y))_{(x,y)\in(\mathbb{Z}_-)^2}$ denotes the kernel of the operator $\varphi(H_k^-)$.

On $\Sigma_{\mathbb{Z}}^\circ$, the spectral measure dN_k^- admits a density with respect to the Lebesgue measure, say n_k^-, and this density is real analytic (see Proposition 5.6).

For $E \in \Sigma_{\mathbb{Z}}^\circ$, define

$$S_k^-(E) := \text{p.v.}\left(\int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E}\right) = \lim_{\varepsilon\to 0+}\left(\int_{-\infty}^{E-\varepsilon} \frac{dN_k^-(\lambda)}{\lambda - E} - \int_{E+\varepsilon}^{+\infty} \frac{dN_k^-(\lambda)}{\lambda - E}\right). \tag{1-3}$$

The existence and analyticity of the Cauchy principal value S_k^- on $\Sigma_{\mathbb{Z}}^\circ$ is guaranteed by the analyticity of n_k^- (see, e.g., [King 2009]). Moreover, for $E \in \Sigma_{\mathbb{Z}}^\circ$, one has

$$S_k^-(E) = \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E - i\varepsilon} - i\pi n_k^-(E). \tag{1-4}$$

In the lower half-plane $\{\operatorname{Im} E < 0\}$, define the function

$$\Xi_k^-(E) := \int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E} + e^{-i \arccos(E/2)} = \int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E} + \frac{E}{2} + \sqrt{\left(\frac{E}{2}\right)^2 - 1}, \tag{1-5}$$

where

- in the first formula, the function $z \mapsto \arccos z$ is the analytic continuation to the lower half-plane of the branch of $\arccos z$ taking values in $[-\pi, 0]$ on the interval $[-1, 1]$;
- in the second formula, the branch of the square root $z \mapsto \sqrt{z^2 - 1}$ has positive imaginary part for $z \in (-1, 1)$.

The function Ξ_k^- is analytic in $\{\operatorname{Im} E < 0\}$ and in a neighborhood of $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$. Moreover, Ξ_k^- vanishes identically if and only if $V \equiv 0$ (see Proposition 5.7).

From now on we assume that $V \not\equiv 0$. In this case, in $\{\operatorname{Im} E < 0\}$ and on $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$, the analytic function Ξ_k^- has only finitely many zeros, each of finite multiplicity (see Proposition 5.7).

We shall need the analogues of the above-defined functions for the already-introduced operator $H_0^+ := H^{\mathbb{N}} = -\Delta + V$ considered on $\ell^2(\mathbb{N})$ with Dirichlet boundary conditions at 0. We define the function N_0^+ as the distribution function of the spectral measure of H_0^+, i.e., for $\varphi \in \mathscr{C}_0^\infty(\mathbb{R})$, we define $\int_{\mathbb{R}} \varphi(\lambda)\, dN_0^+(\lambda) := \varphi(H_0^+)(0,0)$. In the same way as we have defined n_k^-, S_k^- and Ξ_k^- from H_k^-, one can define n_0^+, S_0^+ and Ξ_0^+ from H_0^+. They also satisfy Proposition 5.6, relation (1-4) and Proposition 5.7.

For the description of the resonances, it will be convenient to define the following functions on $\Sigma_{\mathbb{Z}}^\circ$:

$$c^{\mathbb{N}}(E) := i + \frac{\Xi_k^-(E)}{\pi n_k^-(E)} = \frac{1}{\pi n_k^-(E)}(S_k^-(E) + e^{-i\arccos(E/2)}) \tag{1-6}$$

and

$$c^{\mathbb{Z}}(E) := \frac{\dfrac{(S_0^+(E)+e^{-i\arccos(E/2)})(S_k^-(E)+e^{-i\arccos(E/2)})}{n_0^+(E) n_k^-(E)} - \pi^2}{\dfrac{\pi(S_0^+(E)+e^{-i\arccos(E/2)})}{n_0^+(E)} + \dfrac{\pi(S_k^-(E)+e^{-i\arccos(E/2)})}{n_k^-(E)}}. \tag{1-7}$$

We shall see that the zeros of $c^\bullet - i$ play a special role for the resonances of $H_L^\bullet$; therefore, we define

$$\mathfrak{D}^\bullet = \{z \in \Sigma_{\mathbb{Z}}^\circ \mid c^\bullet(z) = i\}. \tag{1-8}$$

The set $\mathfrak{D}$ introduced in Theorem 0.1 is the set $\mathfrak{D}^{\mathbb{Z}} \cap (-2,2)$.

Remark 1.6. Before describing the resonances, let us explain why the operators H_0^+ and H_k^- naturally occur in this study. They respectively are the strong resolvent limits (when $L \to +\infty$ with $L \in p\mathbb{N} + k$) of the operator $H_L^{\mathbb{Z}}$ restricted to $[\![0, L]\!]$ with Dirichlet boundary conditions at 0 and L “seen” from the left- and the right-hand side, respectively.

Indeed, define H_L to be the operator $H_L^{\mathbb{N}}$ restricted to $[\![0, L]\!]$ with Dirichlet boundary conditions at L (see Remark 1.1). Note that H_L is also the operator $H_L^{\mathbb{Z}}$ restricted to $[\![0, L]\!]$ with Dirichlet boundary conditions at 0 and L.

Clearly, the operator H_0^+ is the strong resolvent limit of H_L when $L \to +\infty$.

If $\tilde\tau_L$ denotes the translation by $-L$ that unitarily maps $\ell^2([\![0, L]\!])$ into $\ell^2([\![-L, 0]\!])$, then $\widetilde{H}_L = \tilde\tau_L H_L \tilde\tau_L^*$ converges in the strong resolvent sense to H_k^- when $L \to +\infty$ and $L \equiv k \bmod p$. Indeed, $\tau_L V = \tau_k V$ as V is p-periodic.

1B3. *Description of the resonances closest to the real axis.* Let $(\lambda_l)_{0\le l\le L} = (\lambda_l^L)_{0\le l\le L}$ be the eigenvalues of H_L (that is, the eigenvalues of $H_L^{\mathbb{N}}$ or $H_L^{\mathbb{Z}}$ restricted to $[\![0, L]\!]$ with Dirichlet boundary conditions; see Remark 1.1) listed in increasing order. They are described in Theorem 4.2; those away from the edges of $\Sigma_{\mathbb{Z}}$ are shown to be nicely interspaced points at a distance roughly L^{-1} from one another.

We first state our most general result describing the resonances in a uniform way. We then derive two corollaries describing the behavior of the resonance, first far from the set of exceptional energies $\mathfrak{D}^\bullet$ and second close to an exceptional energy.

Pick a compact interval $I \subset (-2,2) \cap \Sigma_{\mathbb{Z}}^\circ$. For $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$ and $\lambda_l \in I$, for L large, define the complex number

$$\tilde z_l^\bullet = \lambda_l + \frac{1}{\pi n(\lambda_l) L} \cot^{-1} \circ c^\bullet\left[\lambda_l + \frac{1}{\pi n(\lambda_l) L} \cot^{-1} \circ c^\bullet\left(\lambda_l - i\frac{\log L}{L}\right)\right], \tag{1-9}$$

where the branch of $\cot^{-1}$ is the inverse of the branch of $z \mapsto \cot z$ that maps $[0,\pi) \times (0, -\infty)$ onto $\mathbb{C}^+ \setminus \{i\}$.

Note that, by Proposition 5.8, for L sufficiently large we know that, for any l such that $\lambda_l \in I$, one has

$$\operatorname{Im} c^\bullet\left(\lambda_l - i\frac{\log L}{L}\right) \in (0,+\infty)\setminus\{1\}$$

and

$$\operatorname{Im} c^\bullet\left[\lambda_l + \frac{1}{\pi n(\lambda_l)L}\cot^{-1}\circ c^\bullet\left(\lambda_l - i\frac{\log L}{L}\right)\right] \in (0,+\infty)\setminus\{1\}.$$

Thus, the formula (1-9) defines $\tilde{z}_l^\bullet$ properly and in a unique way. Moreover, as the zeros of $E \mapsto c^\bullet(E)-i$ are of finite order, one checks that

$$-\log L \lesssim L\cdot \operatorname{Im}\tilde{z}_l^\bullet \lesssim -1 \quad \text{and} \quad 1 \lesssim L\cdot \operatorname{Re}(\tilde{z}_{l+1}^\bullet - \tilde{z}_l^\bullet), \tag{1-10}$$

where the implicit constants are uniform for l such that $\lambda_l \in I$.

We prove:

Theorem 1.7. *Pick $\bullet \in \{\mathbb{N},\mathbb{Z}\}$ and $k \in \{0,\dots,p-1\}$. Let $E_0 \in (-2,2)\cap\Sigma_{\mathbb{Z}}^\circ$.*

Then there exists $\eta_0 > 0$ and $L_0 > 0$ such that, for $L > L_0$ satisfying $L = k \bmod p$, for each $\lambda_l \in I := [E_0-\eta_0, E_0+\eta_0]$, there exists a unique resonance of $H_L^\bullet$, say $z_l^\bullet$, in the rectangle

$$\left[\tfrac{1}{2}\operatorname{Re}(\tilde{z}_l^\bullet + \tilde{z}_{l-1}^\bullet), \tfrac{1}{2}\operatorname{Re}(\tilde{z}_l^\bullet + \tilde{z}_{l+1}^\bullet)\right] + i[-\eta_0, 0];$$

this resonance is simple and it satisfies $|z_l^\bullet - \tilde{z}_l^\bullet| \lesssim 1/(L\log L)$.

This result calls for a few comments. First, the picture one gets for the resonances can be described as follows (see also Figure 3). As long as λ_l stays away from any zero of $E \mapsto c^\bullet(E)-i$, the resonances are nicely spaced points, as the following proposition proves.

Proposition 1.8. *Pick $\bullet \in \{\mathbb{N},\mathbb{Z}\}$ and $k \in \{0,\dots,p-1\}$. Let $I \subset (-2,2)\cap\Sigma_{\mathbb{Z}}^\circ$ be a compact interval such that $I \cap \mathfrak{D}^\bullet = \varnothing$.*

Then, for L sufficiently large and each $\lambda_l \in I$, the resonance $z_l^\bullet$ admits the expansion

$$z_l^\bullet = \lambda_l + \frac{1}{\pi n(\lambda_l)L}\cot^{-1}\circ c^\bullet(\lambda_l) + O\left(\frac{1}{L^2}\right), \tag{1-11}$$

where the remainder term is uniform in l.

The proof of Proposition 1.8 actually yields a complete asymptotic expansion in powers of L^{-1} for the resonances in this zone (see Section 5B5).

Proposition 1.8 implies Theorem 0.1: we choose $\bullet = \mathbb{Z}$ and $k = 0$, then the set $\mathfrak{D}$ of exceptional points in Theorem 0.1 is exactly $\mathfrak{D}^{\mathbb{Z}} \cap (-2,2)$; to obtain (0-3), it suffices to use the asymptotic form of the Dirichlet eigenvalues given by Theorem 4.2.

Near the zeros of $E \mapsto c^\bullet(E)-i$, the resonances take a "plunge" into the lower half of the complex plane (see Figure 3) and their imaginary part becomes of order $L^{-1}\log L$. Indeed, Theorem 1.7 and (1-9) imply:

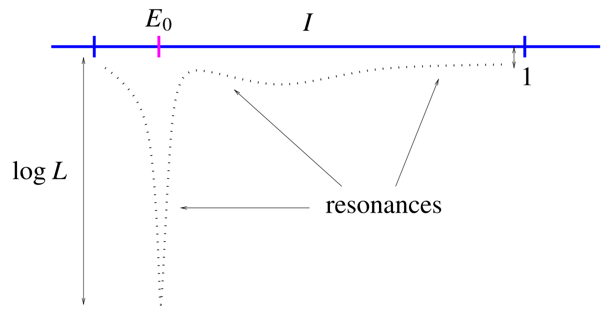


Figure 3. The resonances close to the real axis in the periodic case (after rescaling their imaginary parts by L).

Proposition 1.9. *Pick $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$ and $k \in \{0, \dots, p-1\}$. Let $E_0 \in \mathfrak{D}^\bullet$ be a zero of $E \mapsto c^\bullet(E) - i$ of order q in $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$.*

Then, for $\alpha > 0$ and L sufficiently large, if l is such that $|\lambda_l - E_0| \le L^{-\alpha}$, the resonance $z_l^\bullet$ satisfies

$$\operatorname{Im} z_l^\bullet = \frac{q}{2\pi n(\lambda_l)} \frac{\log\big(|\lambda_l - E_0|^2 + \big(q \log L/(2\pi n(\lambda_l)L)\big)^2\big)}{2L}(1 + o(1)), \tag{1-12}$$

where the remainder term is uniform in l such that $|\lambda_l - E_0| \le L^{-\alpha}$.

When $\bullet = \mathbb{Z}$, the asymptotic (1-12) shows that there can be a “resonance” phenomenon for resonances: when the two functions Ξ_k^- and Ξ_0^+ share a zero at the same real energy, the maximal width of the resonances increases; indeed, the factor in front of $L^{-1} \log L$ is proportional to the multiplicity of the zero of $\Xi_k^- \Xi_0^+$.

1B4. *Description of the low-lying resonances.* The resonances found in Theorem 1.7 are not necessarily the only ones: deeper in the lower complex plane, one may find more resonances. They are related to the zeros of Ξ_k^- when $\bullet = \mathbb{N}$ and of $\Xi_k^- \Xi_0^+$ when $\bullet = \mathbb{Z}$ (see Proposition 5.8).

We now study what happens below the line $\{\operatorname{Im} z = -\eta_0\}$ (see Theorem 1.7) for the resonances of $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$.

The functions Ξ_k^- and Ξ_0^+ are analytic in the lower half-plane and, by Proposition 5.7, they don’t vanish in an neighborhood of $-i\infty$. Hence, the functions Ξ_k^- and Ξ_0^+ have only finitely many zeros in the lower half-plane.

We prove:

Theorem 1.10. *Pick $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$ and $k \in \{0, \dots, p-1\}$. Let $(E_j^\bullet)_{1 \le j \le J}$ be the zeros of $E \mapsto c^\bullet(E) - i$ in $I + i(-\infty, 0)$. Pick $E_0 \in (-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$.*

There exists $\eta_0 > 0$ such that, for $I = E_0 + [-\eta_0, \eta_0]$ and L sufficiently large with $L \equiv k \bmod p$, one has:

- *If $E_0 \notin \{\operatorname{Re} E_j^\bullet \mid 1 \le j \le J\}$, then in the rectangle $I + i(-\infty, 0]$ the only resonances of $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$ are those given by Theorem 1.7.*
- *If $E_0 \in \{\operatorname{Re} E_j^\bullet \mid 1 \le j \le J\}$, then*

- *in the rectangle* $I + i[-\eta_0, 0]$, *the only resonances of* $H_L^{\mathbb{N}}$ *and* $H_L^{\mathbb{Z}}$ *are those given by Theorem 1.7*;
- *in the strip* $I + i[-\infty, -\eta_0]$, *the resonances of* $H_L^\bullet$ *are contained in* $\bigcup_{j=1}^{J} D(E_j^\bullet, e^{-\eta_0 L})$;
- *in* $D(E_j^\bullet, e^{-\eta_0 L})$, *the number of resonances* (*counted with multiplicity*) *is equal to the order of* $E_j^\bullet$ *as a zero of* $E \mapsto c^\bullet(E) - i$.

We see that the total number of resonances below a compact subset of $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$ that do not tend to the real axis when $L \to +\infty$ is finite. These resonances are related to the resonances of H^∞, to which we turn now.

1B5. *The half-line periodic perturbation.* Fix $p \in \mathbb{N}^*$. On $\ell^2(\mathbb{Z})$, we now consider the operator $H^\infty = \Delta + V$, where $V(n) = 0$ for $n \geq 0$ and $V(n + p) = V(n)$ for $n \leq -1$. We prove:

Theorem 1.11. *The resolvent of* H^∞ *can be analytically continued from the upper half-plane through* $(-2, 2) \cap \Sigma_Z^\circ$ *to the lower half-plane. The resulting operator does not have any poles in the lower half-plane or on* $(-2, 2) \cap \Sigma_Z^\circ$.

The resolvent of H^∞ *can be analytically continued from the upper half-plane through* $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ (*resp.* $\Sigma_Z^\circ \setminus [-2, 2]$) *to the lower half-plane*; *the poles of the continuation through* $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ (*resp.* $\Sigma_Z^\circ \setminus [-2, 2]$) *are exactly the zeros of the function* $E \mapsto 1 - e^{i\theta(E)} \int_{\mathbb{R}} 1/(\lambda - E)\, dN_{p-1}^-(\lambda)$ *when continued from the upper half-plane through* $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ (*resp.* $\Sigma_Z^\circ \setminus [-2, 2]$) *to the lower half-plane.*

Remark 1.12. In Theorem 1.11 and below, every time we consider the analytic continuation of a resolvent through some open subset of the real line we implicitly assume the open subset to be nonempty.

In Figure 4, to illustrate Theorem 1.11, assuming that $\Sigma_{\mathbb{Z}}$ (in blue) has a single gap that is contained in $(-2, 2)$, we have drawn the various analytic continuations of the resolvent of H^∞ and the presence or absence of resonances for the different continuations.

Using the same arguments as in the proof of Proposition 5.7, one easily sees that the continuations of the function $E \mapsto 1 - e^{i\theta(E)} \int_{\mathbb{R}} 1/(\lambda - E)\, dN_{p-1}^-(\lambda)$ to the lower half-plane through $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ and $\Sigma_Z^\circ \setminus [-2, 2]$ have at most finitely many zeros and that these zeros are away from the real axis.

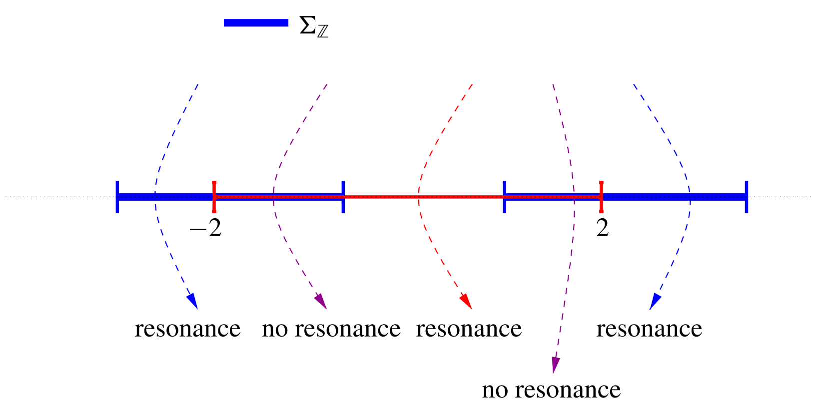


Figure 4. The analytic continuation of the resolvent and resonances for H^∞.

This also implies that the spectrum on H^∞ in $[-2,2]\cup\Sigma_{\mathbb{Z}}$ is purely absolutely continuous except possibly at the points of $\partial\Sigma_{\mathbb{Z}}\cup\{-2,2\}$, where $\partial\Sigma_{\mathbb{Z}}$ is the set of edges of $\Sigma_{\mathbb{Z}}$.

1C. ***The random case.*** We now turn to the random case. Let $V=V_\omega$, where $(V_\omega(n))_{n\in\mathbb{Z}}$ are bounded independent and identically distributed random variables. Assume that the common law of the random variables admits a bounded compactly supported density, say g.

Set $H^{\mathbb{N}}_\omega=-\Delta+V_\omega$ on $\ell^2(\mathbb{N})$ (with Dirichlet boundary condition at 0 for concreteness). Let $\sigma(H^{\mathbb{N}}_\omega)$ be the spectrum of $H^{\mathbb{N}}_\omega$. Consider also $H^{\mathbb{Z}}_\omega=-\Delta+V_\omega$ acting on $\ell^2(\mathbb{Z})$. Then one knows (see, e.g., [Kirsch 2008]) that, ω-almost surely,

$$\sigma(H^{\mathbb{Z}}_\omega)=\Sigma:=[-2,2]+\operatorname{supp} g. \tag{1-13}$$

One has the following description for the spectra $\sigma(H^{\mathbb{N}}_\omega)$ and $\sigma(H^{\mathbb{Z}}_\omega)$:

- ω-almost surely, $\sigma(H^{\mathbb{Z}}_\omega)=\Sigma$; the spectrum is purely punctual; it consists of simple eigenvalues associated to exponentially decaying eigenfunctions (Anderson localization; see, e.g., [Pastur and Figotin 1992; Kirsch 2008]); one can prove that, under the assumptions made above, the whole spectrum is dynamically localized (see, e.g., [Cycon et al. 1987] and references therein).
- For $H^{\mathbb{N}}_\omega$ (see, e.g., [Pastur and Figotin 1992; Carmona and Lacroix 1990]), one has, ω-almost surely, $\sigma(H^{\mathbb{N}}_\omega)=\Sigma\cup K_\omega$, where
  - Σ is the essential spectrum of $H^{\mathbb{N}}_\omega$ and it consists of simple eigenvalues associated to exponentially decaying eigenfunctions;
  - the set K_ω is the discrete spectrum of $H^{\mathbb{N}}_\omega$, which may be empty and depends on ω.

1C1. *The integrated density of states and the Lyapunov exponent.* It is well known (see, e.g., [Pastur and Figotin 1992]) that the integrated density of states of H, say $N(E)$, is defined as the limit

$$N(E)=\lim_{L\to+\infty}\frac{\#\bigl\{\text{eigenvalues of } H^{\mathbb{Z}}_\omega|_{[\![-L,L]\!]} \text{ in } (-\infty,E]\bigr\}}{2L+1}. \tag{1-14}$$

The above limit does not depend on the boundary conditions used to define the restriction $H^{\mathbb{Z}}_\omega|_{[\![-L,L]\!]}$. It defines the distribution function of a probability measure supported on Σ. Under our assumptions on the random potential, N is known to be Lipschitz continuous ([Pastur and Figotin 1992; Kirsch 2008]). Let $n(E)=dN(E)/dE$ be its derivative; it exists for almost all energies. If one assumes more regularity on g, the density of the random variables $(\omega_n)_n$, then the density of states n can be shown to exist everywhere and to be regular (see, e.g., [Cycon et al. 1987]).

One also defines the Lyapunov exponent, say $\rho(E)$, as

$$\rho(E):=\lim_{L\to+\infty}\frac{\log\|T_L(E,\omega)\|}{L+1},$$

where

$$T_L(E;\omega):=\begin{pmatrix}E-V_\omega(L) & -1\\ 1 & 0\end{pmatrix}\times\cdots\times\begin{pmatrix}E-V_\omega(0) & -1\\ 1 & 0\end{pmatrix} \tag{1-15}$$

For any E, ω-almost surely, the Lyapunov exponent is known to exist and to be independent of ω (see, e.g., [Cycon et al. 1987; Pastur and Figotin 1992; Carmona and Lacroix 1990]). It is positive at all energies. Moreover, by the Thouless formula [Cycon et al. 1987], it is continuous for all E and is the harmonic conjugate of $n(E)$.

For $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$, we now define $H^\bullet_{\omega,L}$ to be the operator $-\Delta^\bullet + V_\omega \mathbf{1}_{[\![0,L]\!]}$. The goal of the next sections is to describe the resonances of these operators in the limit $L \to +\infty$.

As in the case of a periodic potential V, the resonances are defined as the poles of the analytic continuation of $z \mapsto (H^\bullet_{\omega,L} - z)^{-1}$ from $\mathbb{C}^+$ through $(-2, 2)$ (see Theorem 1.3).

1C2. *Resonance-free regions.* We again start with a description of the resonance-free region near a compact interval in $(-2, 2)$. As in the periodic case, the size of the $H^\bullet_{\omega,L}$-resonance-free region below a given energy will depend on whether this energy belongs to $\sigma(H^\bullet_\omega)$ or not. We prove:

Theorem 1.13. *Fix $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$. Let I be a compact interval in $(-2, 2)$. Then, ω-a.s., one has:*

(1) *For $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$, if $I \subset \mathbb{R} \setminus \sigma(H^\bullet_\omega)$ then there exists $C > 0$ such that, for L sufficiently large, there are no resonances of $H^\bullet_{\omega,L}$ in the rectangle $\{\operatorname{Re} z \in I,\ 0 \geq \operatorname{Im} z \geq -1/C\}$.*

(2) *If $I \subset \Sigma^\circ$, then for $\varepsilon \in (0, 1)$ there exists $L_0 > 0$ such that, for $L \geq L_0$, there are no resonances of $H^\bullet_{\omega,L}$ in the rectangle $\{\operatorname{Re} z \in I,\ 0 \geq \operatorname{Im} z \geq -e^{-2\eta_\bullet \rho L(1+\varepsilon)})\}$, where*
   - *ρ is the maximum of the Lyapunov exponent $\rho(E)$ on I,*
   - $\eta_\bullet = \begin{cases} 1 & \text{if } \bullet = \mathbb{N}, \\ \frac{1}{2} & \text{if } \bullet = \mathbb{Z}. \end{cases}$

(3) *Pick $v_j = v_j(\omega) \in K_\omega$* (see the description of the spectrum of $H^{\mathbb{N}}_\omega$ just above Section 1C1) *and assume that $\{v_j\} = I^\circ \cap \sigma(H^{\mathbb{N}}_\omega) = I \cap \sigma(H^{\mathbb{N}}_\omega)$ and $I \cap \Sigma = \varnothing$; then there exists $c > 0$ such that, for L sufficiently large, $H^{\mathbb{N}}_{\omega,L}$ has a unique resonance in $\{\operatorname{Re} z \in I,\ -c \leq \operatorname{Im} z \leq 0\}$; moreover, this resonance, say z_j, is simple and satisfies $\operatorname{Im} z_j \asymp -e^{-\rho_j(\omega)L}$ and $|z_j - \lambda_j| \asymp e^{-\rho_j(\omega)L}$ for some $\rho_j(\omega) > 0$ independent of L.*

When comparing point (2) of this result with Theorem 1.5(2), it is striking that the width of the resonance-free region below Σ is much smaller in the random case (it is exponentially small in L) than in the periodic case (it is polynomially small in L). This a consequence of the localized nature of the spectrum, i.e., of the exponential decay of the eigenfunctions of $H^\bullet_\omega$.

1C3. *Description of the resonances closest to the real axis.* We will now see that below the resonance-free strip exhibited in Theorem 1.13 one does find resonances — actually, many of them. We prove:

Theorem 1.14. *Fix $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$. Let I be a compact interval in $(-2, 2) \cap \overset{\circ}{\Sigma}$.*

(1) *For any $\kappa \in (0, 1)$, ω-a.s. one has*
$$\frac{\#\{z \text{ resonance of } H^\bullet_{\omega,L} \mid \operatorname{Re} z \in I,\ 0 > \operatorname{Im} z \geq -e^{-L^\kappa}\}}{L} \to \int_I n(E)\, dE.$$

(2) *For $E \in I$ such that $n(E) > 0$ and $\lambda \in (0, 1)$, define the rectangle*
$$R^\bullet(E, \lambda, L, \varepsilon, \delta) := \left\{ z \in \mathbb{C} \;\middle|\; n(E)|\operatorname{Re} z - E| \leq \tfrac{1}{2}\varepsilon,\ -e^{\eta_\bullet \rho(E)\delta L} \leq e^{2\eta_\bullet \rho(E)\lambda L} \operatorname{Im} z \leq -e^{-\eta_\bullet \rho(E)\delta L} \right\},$$

where $\eta^\bullet$ is as defined in Theorem 1.13; then ω-a.s. one has

$$\lim_{\delta\to 0^+}\lim_{\varepsilon\to 0^+}\lim_{L\to+\infty}\frac{\#\{z \text{ resonances of } H^\bullet_{\omega,L} \text{ in } R^\bullet(E,\lambda,L,\varepsilon,\delta)\}}{L\varepsilon\delta}=1. \tag{1-16}$$

(3) *For $E\in I$ such that $n(E)>0$, define*

$$R^\bullet_\pm(E,1,L,\varepsilon,\delta)=\left\{z\in\mathbb{C}\;\middle|\; n(E)|\mathrm{Re}\,z-E|\le\tfrac{1}{2}\varepsilon,\ -e^{-2\eta_\bullet\rho(E)(1\pm\delta)L}\le\mathrm{Im}\,z<0\right\};$$

then ω-a.s. one has

$$\lim_{\delta\to 0^+}\lim_{\varepsilon\to 0^+}\lim_{L\to+\infty}\frac{\#\{\text{resonances in } R^\bullet_\pm(E,1,L,\varepsilon,\delta)\}}{L\varepsilon\delta}=\begin{cases}1 & \text{if } \pm=-,\\ 0 & \text{if } \pm=+.\end{cases} \tag{1-17}$$

(4) *For $c>0$, ω-a.s. one has*

$$\lim_{L\to+\infty}\frac{\#\{z \text{ resonances of } H^\bullet_{\omega,L} \text{ in } I+i(-\infty,-e^{-2cL}]\}}{L}=\int_I\min\left(\frac{c}{\rho(E)},\,1\right)n(E)\,dE. \tag{1-18}$$

The striking fact is that the resonances are much closer to the real axis than in the periodic case; the lifetime of these resonances is much larger. The resonant states are quite stable, with lifetimes that are exponentially large in the width of the random perturbation. Point (4) is an integral version of point (2). Let us also note here that when $\bullet=\mathbb{Z}$, Theorem 1.14(4) is the statement of Theorem 0.2.

Note that the rectangles $R^\bullet(E,\lambda,L,\varepsilon,\delta)$ are very stretched along the real axis; their side-length in the imaginary part is exponentially small in L whereas their side-length in the real part is of order 1.

To understand Theorem 1.14(2), rescale the resonances of $H^\bullet_{\omega,L}$, say $(z^\bullet_{l,L}(\omega))_l$, as

$$x^\bullet_l=x^\bullet_{l,L}(E,\omega)=n(E)L(\mathrm{Re}\,z^\bullet_{l,L}(\omega)-E)\quad\text{and}\quad y^\bullet_l=y^\bullet_{l,L}(E,\omega)=-\frac{1}{2\eta_\bullet\rho(E)L}\log|\mathrm{Im}\,z^\bullet_{l,L}(\omega)|. \tag{1-19}$$

For $\lambda\in(0,1)$, this rescaling maps the rectangle $R^\bullet(E,\lambda,L,\varepsilon,\delta)$ into $\left\{|x|\le\frac{1}{2}L\varepsilon,\ |y-\lambda|\le\frac{1}{2}\delta\right\}$ and the rectangles $R^\bullet_\pm(E,1,L,\varepsilon,\delta)$ are mapped into $\{|x|\le L\varepsilon/2,\ 1\mp\delta\le y\}$, respectively. The denominator of the quotient in (1-16) is just the area of the rescaled $R^\bullet(E,\lambda,L,\varepsilon,\delta)$ for $\lambda\in(0,1)$ or the rescaled $R^\bullet_+(E,1,L,\varepsilon,\delta)\setminus R^\bullet_-(E,1,L,\varepsilon,0)$. So, (2) states that, in the limit with ε and δ small and L large, the rescaled resonances become uniformly distributed in the rescaled rectangles.

We see that the structure of the set of resonances is very different from the one observed in the periodic case (see Figure 2). We will now zoom in on the resonance even more so as to make this structure clearer. We consider the two-dimensional point process $\xi^\bullet_L(E,\omega)$ defined by

$$\xi^\bullet_L(E,\omega)=\sum_{z^\bullet_{l,L}\text{ resonance of } H^\bullet_{\omega,L}}\delta_{(x^\bullet_l,y^\bullet_l)}, \tag{1-20}$$

where $x^\bullet_l$ and $y^\bullet_l$ are defined by (1-19).

We prove:

Theorem 1.15. *Fix $E\in(-2,2)\cap\Sigma^\circ$ such that $n(E)>0$. Then the point process $\xi^\bullet_L(E,\omega)$ converges weakly to a Poisson process in $\mathbb{R}\times(0,1]$ with intensity 1. That is, for any $p\ge 0$, if $(I_n)_{1\le n\le p}$ (resp.*

$(C_n)_{1\le n\le p}$) *are disjoint intervals of the real line* $\mathbb{R}$ (*resp.* $[0,1]$), *then*

$$\lim_{L\to+\infty}\mathbb{P}\big(\big\{\omega\,\big|\,\#\{j\mid x^\bullet_{l,L}(E,\omega)\in I_n,\ y^\bullet_{l,L}(E,\omega)\in C_n\}=k_n \text{ for } n=1,\dots,p\big\}\big)=\prod_{n=1}^{p}e^{-\mu_n}\frac{(\mu_n)^{k_n}}{k_n!},$$

where $\mu_n:=|I_n||C_n|$ *for* $1\le n\le p$.

This is the analogue of the celebrated result on the Poisson structure of the eigenvalues and localization centers of a random system (see, e.g., [Molchanov 1982; Minami 1996; Germinet and Klopp 2014]).

When considering the model for $\bullet=\mathbb{Z}$, Theorem 1.15 is Theorem 0.3.

In [Klopp 2011], we proved decorrelation estimates that can be used in the present setting to prove:

Theorem 1.16. *Fix* $E\in(-2,2)\cap\Sigma^\circ$ *and* $E'\in(-2,2)\cap\Sigma^\circ$ *such that* $E\neq E'$, $n(E)>0$ *and* $n(E')>0$. *Then the limits of the processes* $\xi^\bullet_L(E,\omega)$ *and* $\xi^\bullet_L(E',\omega)$ *are stochastically independent.*

Due to the rescaling, the above results only give a picture of the resonances in a zone of the type

$$E+L^{-1}[-\varepsilon^{-1},\varepsilon^{-1}]-i[e^{-2\eta_\bullet(1+\varepsilon)\rho(E)L},e^{-2\varepsilon\eta_\bullet\rho(E)L}] \tag{1-21}$$

for $\varepsilon>0$ arbitrarily small.

When L gets large, this rectangle is of a very small width and located very close to the real axis. Theorems 1.14, 1.15 and 1.16 describe the resonances lying closest to the real axis. As a comparison between points (1) and (2) in Theorem 1.14 shows, these resonances are the most numerous.

One can get a number of other statistics (e.g., the distribution of the spacings between the resonances) using the techniques developed for the study of the spectral statistics of a random system in the localized phase (see [Germinet and Klopp 2011; 2014; Klopp 2013]) combined with the analysis developed in Section 6.

1C4. *The description of the low-lying resonances.* It is natural to question what happens deeper in the complex plane. To answer this question, fix an increasing sequence of scales $(\ell_L)_L$ such that

$$\frac{\ell_L}{\log L}\to+\infty \quad\text{as } L\to+\infty \qquad\text{and}\qquad \frac{\ell_L}{L}\to 0\quad\text{as } L\to+\infty. \tag{1-22}$$

We first show that there are only a few resonances below the line $\{\operatorname{Im} z=e^{-\ell_L}\}$, namely:

Theorem 1.17. *Pick* $(\ell_L)_L$ *a sequence of scales satisfying* (1-22) *and* I *as above.*

Then, ω *almost surely, for* L *large one has*

$$\big\{z \text{ resonances of } H^\bullet_{\omega,L} \text{ in } \{\operatorname{Re} z\in I,\ \operatorname{Im} z\le -e^{-\ell_L}\}\big\}=O(\ell_L). \tag{1-23}$$

As we shall show now, after proper rescaling the structure of these resonances is the same as that of the resonances closer to the real axis.

Fix $E\in I$ such that $n(E)>0$. Recall that $(z^\bullet_{l,L}(\omega))_l$ are the resonances of $H_{\omega,L}$. We now rescale the resonances using the sequence $(\ell_L)_L$; this rescaling will select resonances that are further away from the real axis. Define

$$x_l^\bullet = x_{l,\ell_L}^\bullet(\omega) = n(E)\ell_L(\operatorname{Re} z_{l,L}^\bullet(\omega) - E) \quad \text{and} \quad y_j^\bullet = y_{l,\ell_L}^\bullet(\omega) = \frac{1}{2\eta_\bullet \ell_L \rho(E)} \log|\operatorname{Im} z_{l,L}^\bullet(\omega)|. \tag{1-24}$$

Consider now the two-dimensional point process

$$\xi_{L,\ell}^\bullet(E,\omega) = \sum_{z_{l,L}^\bullet \text{ resonance of } H_{\omega,L}^\bullet} \delta_{(x_{l,\ell_L}^\bullet, y_{l,\ell_L}^\bullet)}. \tag{1-25}$$

We prove the following analogue of the results of Theorems 1.14, 1.15 and 1.16 for resonances lying further away from the real axis.

Theorem 1.18. *Fix $E \in (-2,2) \cap \Sigma^\circ$ and $E' \in (-2,2) \cap \Sigma^\circ$ such that $E \neq E'$, $n(E) > 0$ and $n(E') > 0$. Fix a sequence of scales $(\ell_L)_L$ satisfying* (1-22). *Then one has*:

(1) *For $\lambda \in (0,1]$, ω-almost surely,*

$$\lim_{\delta\to 0^+} \lim_{\varepsilon\to 0^+} \lim_{L\to+\infty} \frac{\#\{z \text{ resonances of } H_{\omega,L}^\bullet \text{ in } R^\bullet(E,\lambda,\ell_L,\varepsilon,\delta)\}}{\ell_L \varepsilon\delta} = 1,$$

where $R^\bullet(E,\lambda,L,\varepsilon,\delta)$ is as defined in Theorem 1.14.

(2) *The point processes $\xi_{L,\ell}^\bullet(E,\omega)$ and $\xi_{L,\ell}^\bullet(E',\omega)$ converge weakly to Poisson processes in $\mathbb{R}\times(0,+\infty)$ of intensity* 1.

(3) *The limits of the processes $\xi_{L,\ell}^\bullet(E,\omega)$ and $\xi_{L,\ell}^\bullet(E',\omega)$ are stochastically independent.*

Point (1) shows that, in (1-23), one actually has

$$\left\{z \text{ resonances of } H_{\omega,L}^\bullet \text{ in } \{\operatorname{Re} z \in I,\ \operatorname{Im} z \le -e^{-\ell_L}\}\right\} \asymp \ell_L.$$

Notice also that the effect of the scaling (1-24) is to select resonances that live in the rectangle

$$E + \ell_L^{-1}[-\varepsilon^{-1}, \varepsilon^{-1}] - i[e^{-2\eta_\bullet(1+\varepsilon)\rho(E)\ell_L}, e^{-2\varepsilon\eta_\bullet\rho(E)\ell_L}]$$

This rectangle is now much further away from the real axis than the one considered in Section 1C3.

Modulo rescaling, the picture one gets for resonances in such rectangles is the same we got above in the rectangles (1-21). This description is valid almost all the way from distances to the real axis that are exponentially small in L up to distances that are of order $e^{-(\log L)^\alpha}$, $\alpha > 1$ (see (1-22)).

1C5. *Deep resonances.* One can also study the resonances that are even further away from the real axis in a way similar to what was done in the periodic case in Section 1B4. Define the random potentials on $\mathbb{N}$ and $\mathbb{Z}$

$$\widetilde{V}_{\omega,L}^{\mathbb{N}}(n) = \begin{cases} \omega_{L-n} & \text{for } 0 \le n \le L, \\ 0 & \text{for } L+1 \le n, \end{cases}$$
$$\widetilde{V}_{\omega,\tilde\omega,L}^{\mathbb{Z}}(n) = \begin{cases} 0 & \text{for } n \le -1, \\ \tilde\omega_n & \text{for } 0 \le n \le \left[\tfrac{1}{2}L\right], \\ \omega_{L-n} & \text{for } \left[\tfrac{1}{2}L\right]+1 \le n \le L, \\ 0 & \text{for } L+1 \le n, \end{cases} \tag{1-26}$$

where $\omega=(\omega_n)_{n\in\mathbb{N}}$ and $\tilde\omega=(\tilde\omega_n)_{n\in\mathbb{N}}$ are i.i.d. and satisfy the assumptions of the beginning of Section 1C.

Consider the operators

- $\widetilde{H}^{\mathbb{N}}_{\omega,L}=-\Delta+\widetilde{V}^{\mathbb{N}}_{\omega,L}$ on $\ell^2(\mathbb{N})$ with Dirichlet boundary condition at 0,
- $\widetilde{H}^{\mathbb{Z}}_{\omega,\tilde\omega,L}=-\Delta+\widetilde{V}^{\mathbb{Z}}_{\omega,\tilde\omega,L}$ on $\ell^2(\mathbb{Z})$.

Clearly, the random operator $\widetilde{H}^{\mathbb{N}}_{\omega,L}$ (resp. $\widetilde{H}^{\mathbb{Z}}_{\omega,L}$) has the same distribution as $H^{\mathbb{N}}_{\omega,L}$ (resp. $H^{\mathbb{Z}}_{\omega,L}$). Thus, for the low lying resonances, we are now going to describe those of $\widetilde{H}^{\mathbb{N}}_{\omega,L}$ (resp. $\widetilde{H}^{\mathbb{Z}}_{\omega,L}$) instead of those of $H^{\mathbb{N}}_{\omega,L}$ (resp. $H^{\mathbb{Z}}_{\omega,L}$).

Remark 1.19. The reason for this change of operators is the same as the one why, in the case of the periodic potential, we had to distinguish various auxiliary operators depending on the congruence of L modulo the period p: this gives a meaning to the limiting operators when $L\to+\infty$.

Define the probability measure $dN_\omega(\lambda)$ using its Borel transform by, for $\operatorname{Im} z\neq 0$,

$$\int_{\mathbb{R}}\frac{dN_\omega(\lambda)}{\lambda-z}:=\langle\delta_0,(H^{\mathbb{N}}_\omega-E)^{-1}\delta_0\rangle. \tag{1-27}$$

Consider the function

$$\Xi_\omega(E)=\int_{\mathbb{R}}\frac{dN_\omega(\lambda)}{\lambda-E}+e^{-i\arccos(E/2)}=\int_{\mathbb{R}}\frac{dN_\omega(\lambda)}{\lambda-E}+\tfrac{1}{2}E+\sqrt{\left(\tfrac{1}{2}E\right)^2-1}, \tag{1-28}$$

where the choice of $z\mapsto\arccos z$ and $z\mapsto\sqrt{z^2-1}$ are those described after (1-5).

This random function Ξ_ω is the analogue of Ξ^-_k in the periodic case. One has the analogue of Proposition 5.7:

Proposition 1.20. *If $\omega_0\neq 0$, one has $\Xi_\omega(E)\sim-\omega_0E^{-2}$ as $|E|\to\infty$, $\operatorname{Im}E<0$. Thus, ω-almost surely, Ξ_ω does not vanish identically in $\{\operatorname{Im}E<0\}$.*

Pick $I\subset\Sigma^\circ\cap(-2,2)$ compact. Then, ω-almost surely, the number of zeros of Ξ_ω (counted with multiplicity) in $I+i(-\infty,\varepsilon]$ is asymptotic to $\int_I n(E)/\rho(E)\,dE\,|\log\varepsilon|$ as $\varepsilon\to0^+$; moreover, ω-almost surely, there exists $\varepsilon_\omega>0$ such that all the zeros of Ξ_ω in $I+i[-\varepsilon_\omega,0)$ are simple.

It seems reasonable to believe that, except for the zero at $-i\infty$, ω-almost surely all the zeros of Ξ_ω are simple; we do not prove it.

For the "deep" resonances, we then prove:

Theorem 1.21. *Fix $I\subset\Sigma^\circ\cap(-2,2)$ a compact interval. There exists $c>0$ such that, with probability 1, there exists $c_\omega>0$ such that, for L sufficiently large, one has:*

1. *For each resonance of $\widetilde{H}^{\mathbb{N}}_{\omega,L}$ (resp. $\widetilde{H}^{\mathbb{Z}}_{\omega,\tilde\omega,L}$) in $I+i(-\infty,-e^{-cL}]$, say E, there exists a unique zero of Ξ_ω (resp. $\Xi_\omega\Xi_{\tilde\omega}$), say $\widetilde{E}$, such that $|E-\widetilde{E}|\le e^{-c_\omega L}$.*
2. *Reciprocally, to each zero (counted with multiplicity) of Ξ_ω (resp. $\Xi_\omega\Xi_{\tilde\omega}$) in the rectangle $I+i(-\infty,-e^{-cL}]$, say $\widetilde{E}$, one can associate a unique resonance of $\widetilde{H}^{\mathbb{N}}_{\omega,L}$ (resp. $\widetilde{H}^{\mathbb{Z}}_{\omega,\tilde\omega,L}$), say E, such that $|E-\widetilde{E}|\le e^{-c_\omega L}$.*

One can combine this result with the description of the asymptotic distribution of the resonances given by Theorem 1.18 to obtain the asymptotic distributions of the zeros of the function Ξ_ω near a point $E-i\varepsilon$ when $\varepsilon\to 0^+$. Indeed, let $(z_l(\omega))_l$ be the zeros of Ξ_ω in $\{\operatorname{Im} E<0\}$. Rescale the zeros:

$$x_{l,\varepsilon}(\omega)=n(E)|\log\varepsilon|(\operatorname{Re} z_l(\omega)-E)\quad\text{and}\quad y_{l,\varepsilon}(\omega)=-\frac{1}{2\rho(E)|\log\varepsilon|}\log|\operatorname{Im} z_l(\omega)|;\tag{1-29}$$

and consider the two-dimensional point process $\xi_\varepsilon(E,\omega)$ defined by

$$\xi_\varepsilon(E,\omega)=\sum_{z_l(\omega)\text{ zeros of }\Xi_\omega}\delta_{(x_{l,\varepsilon},y_{l,\varepsilon})}.\tag{1-30}$$

Then one has:

Corollary 1.22. *Fix $E\in I$ such that $n(E)>0$. Then the point process $\xi_\varepsilon(E,\omega)$ converges weakly to a Poisson process in $\mathbb{R}\times\mathbb{R}$ with intensity* 1.

The function Ξ_ω has been studied in [Kunz and Shapiro 2006; 2008], where the average density of its zeros was computed. Here we obtain a more precise result.

1C6. *The half-line random perturbation.* On $\ell^2(\mathbb{Z})$, we now consider the operator $H^\infty_\omega=-\Delta+V_\omega$, where $V_\omega(n)=0$ for $n\geq 0$, $V_\omega(n)=\omega_n$ for $n\leq -1$ and $(\omega_n)_{n\geq 0}$ are i.i.d. and have the same distribution as above. The spectral theory of the continuous analogue of H^∞_ω, i.e., the Schrödinger operator on the real line with a random potential on the half-line, was studied in [Carmona 1983].

Recall that Σ is the almost sure spectrum of $H^{\mathbb{Z}}_\omega$ (on $\ell^2(\mathbb{Z})$). We prove:

Theorem 1.23. *First, ω-almost surely, the resolvent of H^∞_ω does not admit an analytic continuation from the upper half-plane through $(-2,2)\cap\Sigma^\circ$ to any subset of the lower half plane. Nevertheless, ω-almost surely, the spectrum of H^∞_ω in $(-2,2)\cap\Sigma^\circ$ is purely absolutely continuous.*

Second, ω-almost surely, the resolvent of H^∞_ω does admit a meromorphic continuation from the upper half-plane through $(-2,2)\setminus\Sigma$ to the lower half-plane; the poles of this continuation are exactly the zeros of the function $E\mapsto 1-e^{i\theta(E)}\int_{\mathbb{R}}1/(\lambda-E)\,dN_\omega(\lambda)$ when continued from the upper half-plane through $(-2,2)\setminus\Sigma$ to the lower half-plane.

Third, ω-almost surely, the spectrum of H^∞_ω in $\Sigma^\circ\setminus[-2,2]$ is pure point associated to exponentially decaying eigenfunctions; hence, the resolvent of H^∞_ω cannot be continued through $\Sigma^\circ\setminus[-2,2]$.

In Figure 5, to illustrate Theorem 1.23, assuming that $\Sigma_{\mathbb{Z}}$ (in blue) has a single gap that is contained in $(-2,2)$, we have drawn the analytic continuation of the resolvent of H^∞_ω and the associated resonances; we also indicate the real intervals of the spectrum through which the resolvent of H^∞_ω does not admit an analytic continuation and the spectral type of H^∞_ω in the intervals.

Let us also note here that if $0\in\operatorname{supp} g$ (where g is the density of the random variables defining the random potential) then, by (1-13), one has $[-2,2]\subset\Sigma$. In this case, there is no possibility to continue the resolvent of H^∞_ω to the lower half-plane passing through $[-2,2]$.

Comparing Theorem 1.23 to Theorem 1.11, we see that, as for the operator H^∞, when continued through $(-2,2)\cap\Sigma^\circ$ the operator H^∞_ω does not have any resonances, but for very different reasons.

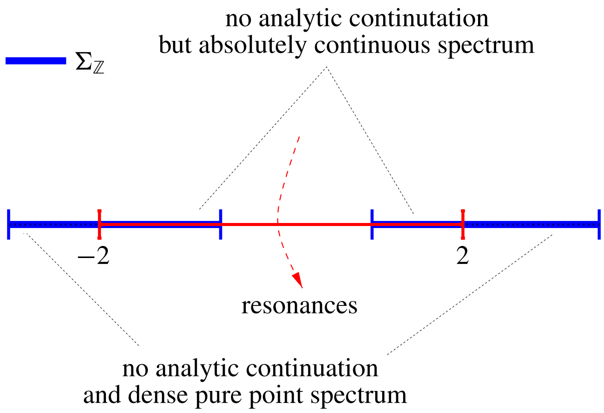


Figure 5. The analytic continuation of the resolvent and resonances for H^{∞}_{ω}.

When one does the continuation through $(-2,2)\setminus\Sigma$, one sees that the number of resonances is finite; "near" the real axis, the continuation of the function $E\mapsto 1-e^{i\theta(E)}\int_{\mathbb{R}} 1/(\lambda-E)\,dN_\omega(\lambda)$ has nontrivial imaginary part and near ∞ it does not vanish.

Theorem 1.23 also shows that the equation studied in [Kunz and Shapiro 2006; 2008], i.e., $\Xi_\omega(E)=0$, does not describe the resonances of H^{∞}_{ω} as is claimed in these papers: these resonances do not exist as there is no analytic continuation of the resolvent of H^{∞}_{ω} through $(-2,2)\cap\Sigma$! As is shown in Theorem 1.21, the solutions to the equation $\Xi_\omega(E)=0$ give an approximation to the resonances of $H^{\mathbb{N}}_{\omega,L}$ (see Theorem 1.21).

1D. ***Outline of and reading guide to the paper.*** In the present section, we shall explain the main ideas leading to the proofs of the results presented above.

In Section 2, we prove Theorem 1.3; this proof is classical. As a consequence of the proof, one sees that, in the case of the half-lattice $\mathbb{N}$ (resp. lattice $\mathbb{Z}$), the resonances are the eigenvalues of a rank-one (resp. rank-two) perturbation of $(-\Delta+V)|_{[\![0,L]\!]}$ with Dirichlet boundary condition. The perturbation depends in an explicit way on the resonance. This yields a closed equation for the resonances in terms of the eigenvalues and normalized eigenfunctions of the Dirichlet restriction $(-\Delta+V)|_{[\![0,L]\!]}$. To obtain a description of the resonances we then are in need of a "precise" description of the eigenvalues and normalized eigenfunctions. Actually, the only information needed on the normalized eigenfunctions is their weight at the point L (and the point 0 in the full lattice case), 0 and L being the endpoints of $[\![0,L]\!]$.

In Section 3, we solve the two equations obtained previously under the condition that the weight of the normalized eigenfunctions at L (and 0) be much smaller than the spacing between the Dirichlet eigenvalues. This condition entails that the resonance equation we want to solve essentially factorizes and become very easy to solve (see Theorems 3.1, 3.2 and 3.3), i.e., it suffices to solve it near any given Dirichlet eigenvalue.

For periodic potentials, the condition that the eigenvalue spacing is much larger than the weight of the normalized eigenfunctions at L (and 0) is not satisfied: both quantities are of the same order of magnitude (see Theorem 4.2) for the Dirichlet eigenvalues in the bulk of the spectrum, i.e., the vast majority of them. This is a consequence of the extended nature of the eigenfunctions in this case. Therefore, we find another way to solve the resonance equation. This way goes through a more precise description of the Dirichlet eigenvalues and normalized eigenfunctions which is the purpose of Theorem 4.2. We use this description to reduce the resonance equation to an effective equation (see Theorem 5.1) up to errors of order $O(L^{-\infty})$. It is important to obtain errors of at most that size. Indeed, the effective equation may have solutions to any order (the order is finite and only depends on V but it is unknown); thus, to obtain solutions to the true equation from solutions to the effective equation with a good precision, one needs the two equations to differ by at most $O(L^{-\infty})$. We then solve the effective equation and, in Section 5B, prove the results of Section 1B.

On the other hand, for random potentials, it is well known that the eigenfunctions of the Dirichlet restriction $(-\Delta + V)|_{[\![0,L]\!]}$ are exponentially localized and, for most of them localized, far from the edge of $[\![0, L]\!]$. Thus, their weight at L (and 0 in the full lattice case) is typically exponentially small in L; the eigenvalue spacing however is typically of order L^{-1}. We can then use the results of Section 3 to solve the resonance equation. The real part of a given resonance is directly related to a Dirichlet eigenvalue and its imaginary part to the weight of the corresponding eigenfunction at L (and 0 in the full lattice case). The main difficulty is to find the asymptotic behavior of this weight. Indeed, while it is known that, in the random case, eigenfunctions decay exponentially away from a localization center and that, for the full random Hamiltonian (i.e., the Hamiltonian on the line or half-line with a random potential), at infinity this decay rate is given by the Lyapunov exponent, to the best of our knowledge, before the present work, it was not known at which length scale this Lyapunov behavior sets in (with a good probability). Answering this question is the purpose of Theorems 6.4 and 6.5 proved in Section 6C: we show that, for the one-dimensional Anderson model, for $\delta > 0$ arbitrary, on a box of size L sufficiently large, all the eigenfunctions exhibit an exponential decay (we obtain both an upper and a lower bound on the eigenfunctions) at a rate equal to the Lyapunov exponent at the corresponding energy (up to an error of size δ) as soon as one is at a distance δL from the corresponding localization center.

These bounds give estimates on the weight of most eigenfunctions at the point L (and 0 in the full lattice case); this is directly related to the distance of the corresponding localization center to the points L (and 0). One can then transform the known results on the statistics of the (rescaled) eigenvalues and (rescaled) localization centers into statistics of the (rescaled) resonances. This is done in Section 6B and proves most of the results in Section 1C.

Finally, Section 6D is devoted to the study of the full line Hamiltonian obtained from the free Hamiltonian on one half-line and a random Hamiltonian on the other half-line; it contains in particular the proof of Theorem 1.23.

2. The analytic continuation of the resolvent

Resonances for Jacobi matrices were considered in various works (see, e.g., [Brown et al. 2005; Iantchenko and Korotyaev 2012] and references therein). For the sake of completeness, we provide an independent proof of Theorem 1.3. It follows standard ideas that were first applied in the continuous setting, i.e., for

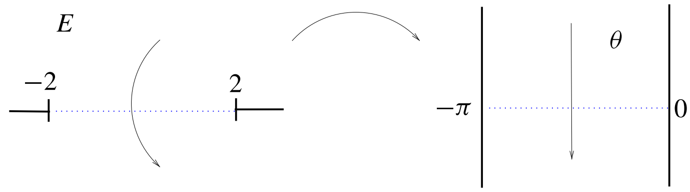


Figure 6. The mapping $E \mapsto \theta(E)$.

partial differential operators instead of finite difference operators (see, e.g., [Sjöstrand and Zworski 1991] and references therein).

The proof relies on the fact that the resolvent of the free Laplace operator can be continued holomorphically from $\mathbb{C}^+$ to $\mathbb{C}\setminus((-\infty,-2]\cup[2,+\infty))$ as an operator valued function from l^2_{comp} to l^2_{loc}. This is an immediate consequence of the fact that, by discrete Fourier transformation, $-\Delta$ is the Fourier multiplier by the function $\theta\mapsto 2\cos\theta$.

Indeed, for $-\Delta$ on $\ell^2(\mathbb{Z})$ and $\operatorname{Im} E>0$, one has, for $(n,m)\in\mathbb{Z}$ (assume $n-m\ge 0$),

$$\begin{aligned}\langle\delta_n,(-\Delta-E)^{-1}\delta_m\rangle &= \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{-i(n-m)\theta}}{2\cos\theta-E}\,d\theta=\frac{1}{2i\pi}\int_{|z|=1}\frac{z^{n-m}}{z^2-Ez+1}\,dz\\ &=\frac{1}{2\sqrt{\left(\frac12 E\right)^2-1}}\left(\tfrac12 E-\sqrt{\left(\tfrac12 E\right)^2-1}\right)^{n-m}=\frac{e^{i(n-m)\theta(E)}}{\sin\theta(E)},\end{aligned} \tag{2-1}$$

where $E=2\cos\theta(E)$ and $\theta=\theta(E)$ is chosen so that $\operatorname{Im}\theta>0$ and $\operatorname{Re}\theta\in(-\pi,0)$ for $\operatorname{Im}E>0$. The choice satisfies $\theta(\bar E)=\overline{\theta(E)}$.

The map $E\mapsto\theta(E)$ can be continued analytically from $\mathbb{C}^+$ to the cut plane $\mathbb{C}\setminus((-\infty,-2]\cup[2,+\infty))$ as shown in Figure 6.

The continuation is one-to-one and onto from $\mathbb{C}\setminus((-\infty,-2]\cup[2,+\infty))$ to $(-\pi,0)+i\mathbb{R}$. It defines a choice of $E\mapsto\arccos\left(\frac12 E\right)=\theta(E)$.

Clearly, using (2-1), this continuation yields an analytic continuation of $R_0^{\mathbb{Z}}:=(-\Delta-E)^{-1}$ from $\{\operatorname{Im}E>0\}$ to $\mathbb{C}\setminus((-\infty,-2]\cup[2,+\infty))$ as an operator from l^2_{comp} to l^2_{loc}.

Let us now turn to the half-line operator, i.e., $-\Delta$ on $\mathbb{N}$ with Dirichlet condition at 0. Pick E such that $\operatorname{Im}E>0$ and set $E=2\cos\theta$, where $\theta=\theta(E)$ is chosen as above. If, for $v\in\mathbb{C}^{\mathbb{N}}$ bounded and $n\ge -1$, one sets $v_{-1}=0$ and

$$[R_0^{\mathbb{N}}(E)(v)]_n=\frac{1}{2i\sin\theta(E)}\sum_{j=-1}^{n}v_j\sin((n-j)\theta(E))-e^{i\theta(E)}\frac{\sin((n+1)\theta(E))}{2i\sin\theta(E)}\sum_{j\ge 0}e^{ij\theta(E)}v_j, \tag{2-2}$$

then, for $\operatorname{Im}E>0$, a direct computations shows that:

(1) For $v\in\ell^2(\mathbb{N})$, the vector $R_0^{\mathbb{N}}(E)(v)$ is in the domain of the Dirichlet Laplacian on $\ell^2(\mathbb{N})$, i.e., $[R_0^{\mathbb{N}}(E)(v)]_{-1}=0$.

(2) For $n \geq 0$, one checks that

$$[R_0^{\mathbb{N}}(E)(v)]_{n+1} + [R_0^{\mathbb{N}}(E)(v)]_{n-1} - E[R_0^{\mathbb{N}}(E)(v)]_n = v_n. \tag{2-3}$$

(3) $R_0^{\mathbb{N}}(E)$ defines a bounded map from $\ell^2(\mathbb{N})$ to itself.

Thus, $R_0^{\mathbb{N}}(E)$ is the resolvent of the Dirichlet Laplacian on $\mathbb{N}$ at energy E for $\operatorname{Im} E > 0$.

Using the continuation of $E \mapsto \theta(E)$, (2-2) yields an analytic continuation of the resolvent $R_0^{\mathbb{N}}(E)$ as an operator from l^2_{comp} to l^2_{loc}.

Remark 2.1. Note that the resolvent $R_0^{\mathbb{N}}(E)$ at an energy E with $\operatorname{Im} E < 0$ is given by (2-2) with $\theta(E)$ replaced by $-\theta(E)$. For (2-2), one has to assume that $(v_j)_{j\in\mathbb{N}}$ decays fast enough at ∞.

To deal with the perturbation V, we proceed in the same way on $\mathbb{Z}$ and on $\mathbb{N}$. Set $V^L = V\mathbf{1}_{[\![0,L]\!]}$ (viewed as a function on $\mathbb{N}$ or $\mathbb{Z}$ depending on the case). Letting $R_0(E)$ be either $R_0^{\mathbb{Z}}(E)$ or $R_0^{\mathbb{N}}(E)$, we compute

$$-\Delta + V^L - E = (-\Delta - E)(1 + R_0(E)V^L) = (1 + V^L R_0(E))(-\Delta_L - E).$$

Thus it suffices to check that the operator $R_0(E)V^L$ (resp. $V^L R_0(E)$) can be analytically continued as an operator from l^2_{loc} to l^2_{loc} (resp. l^2_{comp} to l^2_{comp}). This follows directly from (2-2) and the fact V^L has finite rank.

To complete the proof of Theorem 1.3, we just note that, since

- $E \mapsto R_0(E)V^L$ (resp. $E \mapsto V^L R_0(E)$) is a finite-rank, operator-valued function, analytic on the connected set $\mathbb{C} \setminus ((-\infty, -2] \cup [2, +\infty))$,
- -1 is not an eigenvalue of $R_0(E)V^L$ (resp. $V^L R_0(E)$) for $\operatorname{Im} E > 0$,

by the Fredholm principle, the set of energies E for which -1 is an eigenvalue of $R_0(E)V^L$ (resp. $V^L R_0(E)$) is discrete. Hence, the set of resonances is discrete.

This completes the proof of the first part of Theorem 1.3. To prove the second part, we will first write a characteristic equation for resonances. The bound on the number of resonances will then be obtained through a bound on the number of solutions to this equation.

2A. ***A characteristic equation for resonances.*** In the literature, we did not find a characteristic equation for the resonances in a form suitable for our needs. The characteristic equation we derive will take different forms depending on whether we deal with the half-line or the full line operator. But in both cases, the coefficients of the characteristic equation will be constructed from the spectral data (i.e., the eigenvalues and eigenfunctions) of the operator H_L (see Remark 1.6).

2B. ***In the half-line case.*** We first consider $H_L^{\mathbb{N}}$ on $\ell^2(\mathbb{N})$ and prove:

Theorem 2.2. *Consider the operator H_L defined as $H_L^{\mathbb{N}}$ restricted to $[\![0, L]\!]$ with Dirichlet boundary conditions at L and define*:

- $(\lambda_j)_{0\leq j\leq L} = (\lambda_j(L))_{0\leq j\leq L}$ *are the Dirichlet eigenvalues of $H_L^{\mathbb{N}}$ ordered so that $\lambda_j < \lambda_{j+1}$.*
- $a_j^{\mathbb{N}} = a_j^{\mathbb{N}}(L) = |\varphi_j(L)|^2$, *where $\varphi_j = (\varphi_j(n))_{0\leq n\leq L}$ is a normalized eigenvector associated to λ_j.*

Then an energy E is a resonance of $H_L^{\mathbb{N}}$ if and only if

$$S_L(E) := \sum_{j=0}^{L} \frac{a_j^{\mathbb{N}}}{\lambda_j - E} = -e^{-i\theta(E)}, \quad E = 2\cos\theta(E), \tag{2-4}$$

$\theta(E)$ being chosen so that $\operatorname{Im}\theta(E) > 0$ *and* $\operatorname{Re}\theta(E) \in (-\pi, 0)$ *when* $\operatorname{Im} E > 0$.

Let us note that

$$a_j^{\mathbb{N}}(L) > 0 \quad \text{for all } 0 \le j \le L \qquad \text{and} \qquad \sum_{j=0}^{L} a_j^{\mathbb{N}}(L) = \sum_{j=0}^{L} |\varphi_j(L)|^2 = 1. \tag{2-5}$$

Proof of Theorem 2.2. By the proof of the first statement of Theorem 1.3 (see the beginning of Section 2), we know that an energy E is a resonance if and only if -1 if an eigenvalue of $R_0(E)V^L$, where $R_0(E)$ is defined by (2-2). Pick E an resonance and let $u = (u_n)_{n\ge 0}$ be a resonant state that is an eigenvector of $R_0(E)V^L$ associated to the eigenvalue -1. As $V_n^L = 0$ for $n \ge L+1$, (2-2) yields that, for $n \ge L+1$, $u_n = \beta e^{in\theta(E)}$ for some fixed $\beta \in \mathbb{C}^*$. As $u = -R_0(E)V^L u$, for $n \ge L+1$ it satisfies $u_{n+1} + u_{n-1} = Eu_n$. Thus, $u_{L+1} = e^{i\theta(E)}u_L$ and, by (2-3), u is a solution to the eigenvalues problem

$$\begin{cases} u_{n+1} + u_{n-1} + V_n u_n = E u_n & \text{for all } n \in [\![0, L]\!], \\ u_{-1} = 0, \\ u_{L+1} = e^{i\theta(E)} u_L. \end{cases}$$

This can be equivalently be rewritten as

$$\begin{pmatrix} V_0 & 1 & 0 & \cdots & & 0 \\ 1 & V_1 & 1 & 0 & & \\ \vdots & \ddots & \ddots & \ddots & & \\ & 0 & 1 & V_{L-1} & & 1 \\ 0 & \cdots & 0 & 1 & & V_L + e^{i\theta(E)} \end{pmatrix} \begin{pmatrix} u_0 \\ \\ \vdots \\ \\ u_L \end{pmatrix} = E \begin{pmatrix} u_0 \\ \\ \vdots \\ \\ u_L \end{pmatrix}. \tag{2-6}$$

The matrix in (2-6) is the Dirichlet restriction of $H_L^{\mathbb{N}}$ to $[\![0, L]\!]$ perturbed by the rank-one operator $e^{i\theta(E)}\delta_L \otimes \delta_L$. Thus, by rank-one perturbation theory (see, e.g., [Simon 1995]), an energy E is a resonance if and only if satisfies (2-4).

This completes the proof of Theorem 2.2. □

Proof of Theorem 1.3. Let us now complete the proof of Theorem 1.3 for the operator on the half-line. Let us first note that, for $\operatorname{Im} E > 0$, the imaginary part of the left-hand side of (2-4) is positive by (2-7). On the other hand, the imaginary part of the right-hand side of (2-4) is equal to $-e^{\operatorname{Im}\theta(E)}\sin(\operatorname{Re}\theta(E))$ and, thus, is negative (recall that $\operatorname{Re}\theta(E) \in (-\pi, 0)$ (see Figure 1). Thus, as already emphasized, (2-4) has no solution in the upper half-plane or on the interval $(-2, 2)$.

Clearly, (2-4) is equivalent to the polynomial equation of degree $2L+2$ in the variable $z = e^{-i\theta(E)}$

$$\prod_{k=0}^{L}(z^2 - 2\lambda_k z + 1) - \sum_{j=0}^{L} a_j^{\mathbb{N}} \prod_{\substack{0\le k\le L \\ k\ne j}} (z^2 - 2\lambda_k z + 1) = 0. \tag{2-7}$$

We are looking for the solutions to (2-7) in the upper half-plane. As the polynomial in the right-hand side of (2-7) has real coefficients, its zeros are symmetric with respect to the real axis. Moreover, one notices that, by (2-5), 0 is a solution to (2-7). Hence, the number of solutions to (2-7) in the upper half-plane is bounded by L. This completes the proof of Theorem 1.3. □

2C. ***On the whole line.*** Now consider $H_L^{\mathbb{Z}}$ on $\ell^2(\mathbb{Z})$. We prove:

Theorem 2.3. *Using the notations of Theorem 2.2, an energy E is a resonance of $H_L^{\mathbb{Z}}$ if and only if*

$$\det\left(\sum_{j=0}^{L}\frac{1}{\lambda_j-E}\begin{pmatrix}|\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2\end{pmatrix}+e^{-i\theta(E)}\right)=0, \tag{2-8}$$

where $\det(\cdot)$ denotes the determinant of a square matrix, $E=2\cos\theta(E)$ and $\theta(E)$ is chosen as in Theorem 2.2.

So, an energy E is a resonance of $H_L^{\mathbb{Z}}$ if and only if $-e^{-i\theta(E)}$ belongs to the spectrum of the 2×2 matrix

$$\Gamma_L(E):=\sum_{j=0}^{L}\frac{1}{\lambda_j-E}\begin{pmatrix}|\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2\end{pmatrix}. \tag{2-9}$$

Proof of Theorem 2.3. The proof is the same as that of Theorem 2.2 except that now E is a resonance if there exists a nontrivial solution u to the eigenvalues problem

$$\begin{cases}u_{n+1}+u_{n-1}+V_nu_n=Eu_n & \text{for all } n\in[\![0,L]\!],\\ u_{-1}=e^{i\theta(E)}u_0\\ u_{L+1}=e^{i\theta(E)}u_L.\end{cases}$$

This can equivalently be rewritten as

$$\begin{pmatrix}V_0+e^{i\theta(E)} & 1 & 0 & \cdots & 0\\ 1 & V_1 & 1 & 0 & \\ \vdots & \ddots & \ddots & \ddots & \\ & 0 & 1 & V_{L-1} & 1\\ 0 & \cdots & 0 & 1 & V_L+e^{i\theta(E)}\end{pmatrix}\begin{pmatrix}u_0\\ \\ \vdots\\ \\ u_L\end{pmatrix}=E\begin{pmatrix}u_0\\ \\ \vdots\\ \\ u_L\end{pmatrix}.$$

Thus, using rank-one perturbations twice, we find that an energy E is a resonance if and only if

$$\left(1+e^{i\theta(E)}\sum_{j=0}^{L}\frac{|\varphi_j(0)|^2}{\lambda_j-E}\right)\left(1+e^{i\theta(E)}\sum_{j=0}^{L}\frac{|\varphi_j(L)|^2}{\lambda_j-E}\right)=e^{2i\theta(Ek)}\sum_{0\le j,j'\le L}\frac{\varphi_j(L)\varphi_{j'}(0)\overline{\varphi_{j'}(L)\varphi_j(0)}}{(\lambda_j-E)(\lambda_{j'}-E)},$$

that is, if and only if (2-8) holds. This completes the proof of Theorem 2.3. □

Let us now complete the proof of Theorem 1.3 for the operator on the full line. Let us first show that (2-8) has no solution in the upper half-plane. If $-e^{-i\theta(E)}$ belongs to the spectrum of the matrix defined by (2-8) and $u \in \mathbb{C}^2$ is a normalized eigenvector associated to $-e^{-i\theta(E)}$, one has

$$\sum_{j=0}^{L} \frac{1}{\lambda_j - E} \left| \left\langle \begin{pmatrix} \varphi_j(L) \\ \varphi_j(0) \end{pmatrix}, u \right\rangle \right|^2 = -e^{-i\theta(E)}.$$

This is impossible in the upper half-plane and on $(-2, 2)$ as the two sides of the equation have imaginary parts of opposite signs.

Note that

$$\sum_{j=0}^{L} \begin{pmatrix} \varphi_j(L) \\ \varphi_j(0) \end{pmatrix} \overline{(\varphi_j(L) \;\; \varphi_j(0))} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note also that $-e^{-i\theta(E)}$ is an eigenvalue of (2-8) if and only if it satisfies

$$1 + e^{i\theta(E)} \sum_{j=0}^{L} \frac{|\varphi_j(L)|^2 + |\varphi_j(0)|^2}{\lambda_j - E} = -\tfrac{1}{2} e^{2i\theta(E)} \sum_{0 \le j, j' \le L} \frac{1}{(\lambda_j - E)(\lambda_{j'} - E)} \begin{vmatrix} \varphi_j(0) & \varphi_{j'}(0) \\ \varphi_j(L) & \varphi_{j'}(L) \end{vmatrix}^2. \quad (2\text{-}10)$$

As the eigenvalues of H_L are simple, one computes

$$\sum_{0 \le j, j' \le L} \frac{1}{(\lambda_j - E)(\lambda_{j'} - E)} \begin{vmatrix} \varphi_j(0) & \varphi_{j'}(0) \\ \varphi_j(L) & \varphi_{j'}(L) \end{vmatrix}^2 = 2 \sum_{0 \le j \le L} \frac{1}{\lambda_j - E} \sum_{j' \ne j} \frac{1}{\lambda_{j'} - \lambda_j} \begin{vmatrix} \varphi_j(0) & \varphi_{j'}(0) \\ \varphi_j(L) & \varphi_{j'}(L) \end{vmatrix}^2. \quad (2\text{-}11)$$

Thus, (2-10) is equivalent to the polynomial equation of degree $2(L+1)$ in the variable $z = e^{-i\theta(E)}$

$$z \prod_{k=0}^{L} (z^2 - \lambda_k z + 1) - \sum_{j=0}^{L} (2a_j^{\mathbb{Z}} z + b_j^{\mathbb{Z}}) \prod_{\substack{0 \le k \le L \\ k \ne j}} (z^2 - \lambda_k z + 1) = 0, \quad (2\text{-}12)$$

where we have defined

$$a_j^{\mathbb{Z}} := \frac{1}{2}(|\varphi_j(L)|^2 + |\varphi_j(0)|^2) = \frac{1}{2} \left\| \begin{pmatrix} \varphi_j(L) \\ \varphi_j(0) \end{pmatrix} \right\|^2 = \frac{1}{2} \left\| \begin{pmatrix} |\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L) \\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2 \end{pmatrix} \right\| \quad (2\text{-}13)$$

and

$$b_j^{\mathbb{Z}} := \sum_{j' \ne j} \frac{1}{\lambda_{j'} - \lambda_j} \begin{vmatrix} \varphi_j(0) & \varphi_{j'}(0) \\ \varphi_j(L) & \varphi_{j'}(L) \end{vmatrix}^2.$$

The sequence $(a_j^{\mathbb{Z}})_j$ also satisfies (2-5). Taking $|E|$ to $+\infty$ in (2-11), one notes that

$$\sum_{j=0}^{L} b_j^{\mathbb{Z}} = 0 \quad \text{and} \quad \sum_{j=0}^{L} \lambda_j b_j^{\mathbb{Z}} = -\frac{1}{2} \sum_{0 \le j, j' \le L} \begin{vmatrix} \varphi_j(0) & \varphi_{j'}(0) \\ \varphi_j(L) & \varphi_{j'}(L) \end{vmatrix}^2 = -1. \quad (2\text{-}14)$$

We are looking for the solutions to (2-12) in the upper half-plane. As the polynomial in the right-hand side of (2-12) has real coefficients, its zeros are symmetric with respect to the real axis. Moreover, one notices that, by (2-14), 0 is a root of order two of the polynomial in (2-12). Hence, as the polynomial has degree $2L + 3$, the number of solutions to (2-12) in the upper half-plane is bounded by L. This completes the proof of Theorem 1.3.

3. General estimates on resonances

By Theorems 2.2 and 2.3, we want to solve equations (2-4) and (2-8) in the lower half-plane. We first derive some general estimates for zones in the lower half-plane free of solutions to equations (2-4) and (2-8) (i.e., resonant-free zones for the operators $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$) and then a result on the existence of solutions to equations (2-4) and (2-8) (i.e., resonances for the operators $H_L^{\mathbb{N}}$ and $H_L^{\mathbb{Z}}$).

3A. ***General estimates for resonant-free regions.*** We keep the notations of Theorems 2.2 and 2.3. To simplify the notations in the theorems of this section, we will write a_j for either $a_j^{\mathbb{N}}$ when solving (2-4) or $a_j^{\mathbb{Z}}$ when solving (2-8). We will specify the superscript only when there is risk of confusion.

We first prove:

Theorem 3.1. *Fix $\delta > 0$. Then there exists $C > 0$ (independent of V and L) such that, for any L and $j \in \{0, \dots, L\}$ with $-4+\delta \le \lambda_{j-1}+\lambda_j < \lambda_{j+1}+\lambda_j \le 4-\delta$, equations* (2-4) *and* (2-8) *have no solution in the set* (see Figure 7)

$$U_j := \left\{E \in \mathbb{C} \,\middle|\, \operatorname{Re} E \in \left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right],\ 0 \ge C \cdot \theta'_\delta \operatorname{Im} E > -a_j d_j^2 |\sin \operatorname{Re} \theta(E)|\right\}, \quad (3\text{-}1)$$

where the map $E \mapsto \theta(E)$ is as defined in Section 2 and we have set

$$d_j := \min(\lambda_{j+1}-\lambda_j,\ \lambda_j-\lambda_{j-1},\ 1) \quad \textit{and} \quad \theta'_\delta := \max_{|E|\le 2-\delta/2} |\theta'(E)|. \quad (3\text{-}2)$$

In Theorem 3.1 there are no conditions on the numbers $(a_j)_j$ or $(d_j)_j$ except their being positive. In our application to resonances, this holds. Theorem 3.1 becomes optimal when $a_j \ll d_j^2$. In our application to resonances, for periodic operators one has $a_j \asymp L^{-1}$ and $d_j \asymp L^{-1}$ (see Theorem 5.2), and for random operators one has $a_j \asymp e^{-cL}$ and $d_j \gtrsim L^{-4}$ (see Theorem 6.4 and (6-10)). Thus, in the random case Theorem 3.1 will provide an optimal strip free of resonances, whereas in the periodic case we will use a much more precise computation (see Theorem 5.1) to obtain sharp results.

When $a_j \ll d_j^2$, one proves the existence of another resonant-free region near a energy λ_j, namely:

Theorem 3.2. *Fix $\delta > 0$. Pick $j \in \{0, \dots, L\}$ such that $-4+\delta < \lambda_{j-1}+\lambda_j < \lambda_{j+1}+\lambda_j < 4-\delta$. There exists $C > 0$ (depending only on δ) such that, for any L, if $a_j \le d_j^2/C^2$ then equations* (2-4) *and* (2-8) *have no solution in the set* (see Figure 7)

$$\begin{aligned}\widetilde{U}_j := &\left\{E \in \mathbb{C} \,\middle|\, \operatorname{Re} E \in \left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \lambda_j - C a_j\right] \cup \left[\lambda_j + C a_j, \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right],\ -C a_j \le \operatorname{Im} E \le -\frac{a_j d_j^2}{C}\right\} \\ &\cup \left\{E \in \mathbb{C} \,\middle|\, \operatorname{Re} E \in \left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right],\ -\frac{d_j^2}{C} \le \operatorname{Im} E \le -C a_j\right\}. \quad (3\text{-}3)\end{aligned}$$

Theorem 3.2 becomes optimal when a_j is small and d_j is of order one. This will be sufficient to deal with the isolated eigenvalues for both the periodic and the random potential. It will also be sufficient to give a sharp description of the resonant-free region for random potentials. For the periodic potential, we will rely on much more precise computations (see Theorem 5.1).

Note that Theorem 3.2 guarantees that, if d_j is not too small, outside R_j (see Theorem 3.3) resonances are quite far below the real axis.

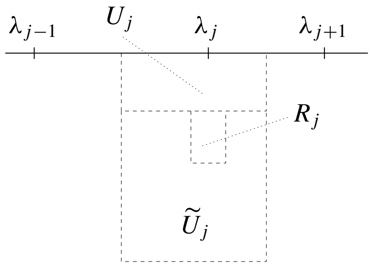


Figure 7. The resonance-free zones U_j and $\widetilde{U}_j$.

Proof of Theorem 3.1. The basic idea of the proof is that, for E close to λ_j, $S_L(E)$ and the matrix $\Gamma_L(E)$ are either large or have a very small imaginary part while, as $-4 < \lambda_{j-1}+\lambda_j < \lambda_{j+1}+\lambda_j < 4$, $e^{-i\theta(E)}$ has a large imaginary part. Thus, (2-4) and (2-8) have no solution in this region.

We start with (2-4). Pick $E \in U_j$ for some C large to be chosen later on. Assume first that $|E-\lambda_j| \le a_j d_j (2+C_0 d_j)^{-1}$ for $C_0 := 2e^{1/C}$. Recall that $0 < a_j, d_j \le 1$. Note that, for C sufficiently large, for $E \in U_j$, one has

$$|\operatorname{Im} e^{-i\theta(E)}| = e^{\operatorname{Im}\theta(E)} |\sin \operatorname{Re}\theta(E)| = e^{\operatorname{Im}[\theta(E)-\theta(\operatorname{Re} E)]} |\sin \operatorname{Re}\theta(E)| \\ \ge e^{\theta'_\delta \operatorname{Im} E} |\sin \operatorname{Re}\theta(E)| \ge e^{-1/C} |\sin \operatorname{Re}\theta(E)| \tag{3-4}$$

and

$$|e^{-i\theta(E)}| \le 1 \le e^{1/C}. \tag{3-5}$$

One estimates

$$|S_L(E)| \ge \frac{a_j}{|\lambda_j - E|} - \sum_{k\ne j} \frac{a_k}{|\lambda_k - E|} \ge \frac{2}{d_j} + C_0 - \sum_{k\ne j} \frac{2a_k}{\min_{k\ne j} |\lambda_k - \lambda_j|} \ge C_0 = 2e^{1/C}. \tag{3-6}$$

Thus, comparing (3-6) and (3-5), we see that (2-4) has no solution in $U_j \cap \{|E-\lambda_j| \le a_j d_j (2+Cd_j)^{-1}\}$.

Assume now that $|E-\lambda_j| > a_j d_j (2+C_0 d_j)^{-1}$. Then, for $E \in U_j$, one has

$$|\operatorname{Im} E| \le \frac{1}{\theta'_\delta C} a_j d_j^2 |\sin \operatorname{Re}\theta(E)|. \tag{3-7}$$

Thus, for $E \in U_j \cap \{|E-\lambda_j| > a_j d_j (2+C_0 d_j)^{-1}\}$, one computes

$$|\operatorname{Im} S_L(E)| \le |\operatorname{Im} E| \left(\frac{a_j}{|\lambda_j - \operatorname{Re} E|^2 + |\operatorname{Im} E|^2} + \frac{4}{d_j^2 + |\operatorname{Im} E|^2} \right) \\ \le \frac{1}{\theta'_\delta C} a_j d_j^2 |\sin(\operatorname{Re}\theta(E))| \left(\frac{(2+C_0 d_j)^2 a_j}{a_j^2 d_j^2} + \frac{4}{d_j^2} \right) \\ \le \frac{4}{\theta'_\delta C} (1+e^{1/C})^2 |\sin(\operatorname{Re}\theta(E))| \le \frac{1}{2} e^{-1/C} |\sin(\operatorname{Re}\theta(E))| \tag{3-8}$$

provided C satisfies $8e^{1/C}(1+e^{1/C})^2 < \theta'_\delta C$.

Hence, the comparison of (3-4) with (3-8) shows that (2-4) has no solution in

$$U_j \cap \{|E-\lambda_j| > a_j d_j (2+C_0 d_j)^{-1}\}$$

if we choose C large enough (independent of $(a_j)_j$ and $(\lambda_j)_j$). Thus, we have proved that, for some $C>0$ large enough (independent of $(a_j)_j$ and $(\lambda_j)_j$), (2-4) has no solution in U_j.

Let us now turn to the case of (2-8). The basic ideas are the same as for (2-4). Consider the matrix $\Gamma_L(E)$ defined by (2-9). The summands in (2-9) are hermitian, of rank 1, and their norm is given by (2-13).

Assume that $E \in U_j$ is a solution to (2-8). Define the vectors

$$v_j := a_j^{-1/2}\begin{pmatrix}\varphi_j(L)\\ \varphi_j(0)\end{pmatrix} \quad \text{for } j \in \{0,\dots,L\}.$$

Here, $a_j = a_j^{\mathbb{Z}}$.

Note that, by definition of a_j, one has $\|v_j\|^2 = 2$. Pick u in C^2 a normalized eigenvector of $\Gamma_L(E)$ associated to the eigenvalue $-e^{-i\theta(E)}$. Thus, u satisfies

$$\sum_{j=0}^{L} \frac{a_j\langle v_j, u\rangle v_j}{\lambda_j - E} = -e^{-i\theta(E)}u. \tag{3-9}$$

Note that, by assumption, one has

$$\sup_{E\in U_j}\left\|\sum_{k\neq j}\frac{a_k\langle v_k,u\rangle v_k}{\lambda_k - E}\right\| \lesssim \frac{1}{d_j} \quad\text{and}\quad \left|\operatorname{Im}\left(\sum_{k\neq j}\frac{a_k|\langle v_k,u\rangle|^2}{\lambda_k-E}\right)\right| \lesssim \frac{|\operatorname{Im} E|}{d_j^2}, \tag{3-10}$$

where the constants are independent of C, the one defining U_j.

Taking the (real) scalar product of (3-9) with $\bar{u}$, and then the imaginary part, we obtain

$$-\frac{a_j|\langle v_j,u\rangle|^2 \operatorname{Im} E}{|\lambda_j - E|^2} + \operatorname{Im}(e^{-i\theta(E)}) = O\left(\frac{|\operatorname{Im} E|}{d_j^2}\right).$$

Thus, for $E \in U_j$, as $a_j \le 1$, for C in (3-1) sufficiently large (depending only on δ),

$$\frac{a_j|\langle v_j,u\rangle|^2|\operatorname{Im} E|}{|\lambda_j - E|^2} \ge \frac{1}{2}|\sin \operatorname{Re}\theta(E)|.$$

Hence, for a solution to (2-8) in U_j and u as above, one has

$$|\lambda_j - E| \le |\langle v_j,u\rangle|\sqrt{\frac{2a_j|\operatorname{Im} E|}{|\sin\operatorname{Re}\theta(E)|}} \le 2\sqrt{\frac{a_j|\operatorname{Im} E|}{|\sin\operatorname{Re}\theta(E)|}}.$$

Hence, by the definition of U_j, for C large we get

$$\left|\frac{a_j}{\lambda_j - E}\right| \ge \frac{C\theta'_\delta}{d_j} \gg \frac{1}{d_j}. \tag{3-11}$$

By (3-10), the operator $\Gamma_L(E)$ can be written as

$$\Gamma_L(E)=\frac{a_j}{\lambda_j-E}v_j\otimes v_j+R_j(E)+iI_j(E), \tag{3-12}$$

where $R_j(E)$ and $I_j(E)$ are selfadjoint (I_j is nonnegative) and satisfy

$$\|R_j(E)\|\lesssim\frac{1}{d_j}\quad\text{and}\quad\|I_j(E)\|\lesssim\frac{|\operatorname{Im}E|}{d_j^2}. \tag{3-13}$$

An explicit computation shows that the eigenvalues of the two-by-two matrix $\frac{a_j}{\lambda_j-E}v_j\otimes v_j+R_j(E)$ satisfy

$$\lambda=\frac{a_j}{\lambda_j-E}\left(1+O\left(\frac{d_j}{C\theta'_\delta}\right)\right)\quad\text{or}\quad|\operatorname{Im}\lambda|\lesssim\frac{|\operatorname{Im}E|}{a_j},$$

where the implicit constants are independent of the one defining U_j.

Thus, by (3-12), using (3-11) and the second estimate in (3-13), we see that the eigenvalues of the matrix $\Gamma_L(E)$ satisfy

$$\lambda=\frac{a_j}{\lambda_j-E}\left(1+O\left(\frac{d_j}{C\theta'_\delta}\right)\right)\quad\text{or}\quad|\operatorname{Im}\lambda|\le\frac{2}{C\theta'_\delta}.$$

Clearly, for C large, no such value can be equal to $-e^{-i\theta(E)}$, being too large — by (3-11) — in the first case or having too small imaginary part in the second. The proof of Theorem 3.1 is complete. □

Proof of Theorem 3.2. Again, we start with the solutions to (2-4). For $z\in\widetilde{U}_j$, we compute

$$\operatorname{Im}S_L(E)=\sum_{k=0}^{L}\frac{a_k\operatorname{Im}E}{(\lambda_k-\operatorname{Re}E)^2+\operatorname{Im}^2E}=\frac{a_j\operatorname{Im}E}{(\lambda_j-\operatorname{Re}E)^2+\operatorname{Im}^2E}+\sum_{\substack{0\le k\le L\\ k\neq j}}\frac{-a_k\operatorname{Im}E}{(\lambda_k-\operatorname{Re}E)^2+\operatorname{Im}^2E}. \tag{3-14}$$

When $-d_j^2/C\le\operatorname{Im}E\le-Ca_j$, the second equality above and (2-5) yield, for C sufficiently large,

$$0\le\operatorname{Im}S_L(E)\lesssim\frac{a_j}{|\operatorname{Im}E|}+\frac{|\operatorname{Im}E|}{d_j^2+\operatorname{Im}^2E}\le\frac{2}{C}. \tag{3-15}$$

On the other hand, for some $K>0$, one has

$$|\operatorname{Im}e^{-i\theta(E)}|\ge|\operatorname{Im}e^{-i\theta(\operatorname{Re}E)}|-\frac{Kd_j^2}{C}.$$

Now, since under the assumptions of Theorem 3.2 one has

$$\min_{E\in[(\lambda_j+\lambda_{j-1})/2,(\lambda_j+\lambda_{j+1})/2]}|\operatorname{Im}e^{-i\theta(E)}|\ge\tfrac14\min\left(\sqrt{16-(\lambda_j+\lambda_{j-1})^2},\ \sqrt{16-(\lambda_j+\lambda_{j+1})^2}\right), \tag{3-16}$$

we obtain that (2-4) has no solution in $\widetilde{U}_j\cap\{-d_j/C\le\operatorname{Im}E\le-Ca_j\}$.

Now pick $E\in\widetilde{U}_j$ such that $-Ca_j\le\operatorname{Im}E\le-a_jd_j^2/C$. Then (3-5) and (2-5) yield, for C sufficiently large,

$$\operatorname{Im}S_L(E)\lesssim\frac{a_j\operatorname{Im}E}{C^2a_j^2+\operatorname{Im}^2E}+\frac{Ca_j}{d_j^2}\le\frac{1}{C}+\frac{1}{2C}.$$

The imaginary part of $e^{-i\theta(E)}$ is estimated as above. Thus, for C sufficiently large, (2-4) has no solution in $\widetilde{U}_j \cap \{-Ca_j \le \operatorname{Im} E \le -a_j d_j^2/C\}$.

The case of (2-8) is studied in exactly the same way except that, as in the proof of Theorem 3.1, one has to replace the study of $S_L(E)$ by that of $\langle \Gamma_L(E)u, u\rangle$ for u a normalized eigenvector of $\Gamma_L(E)$ associated to $-e^{-i\theta(E)}$ and, thus, the coefficient a_k in (3-14) gets multiplied by a factor $|\langle v_k, u\rangle|^2$ that is bounded by 2.

This completes the proof of Theorem 3.2. □

3B. ***The resonances near an “isolated” eigenvalue.*** We will now solve (2-4) near a given λ_j under the additional assumptions that $a_j \ll d_j^2$. By Theorems 3.1 and 3.2, we will do so in the rectangle R_j (see Theorem 3.3 and Figure 7). Actually, we prove that in R_j there is exactly one resonance and give an asymptotic for this resonance in terms of a_j, d_j and λ_j. This result is going to be applied to the case of random V and to that of isolated eigenvalues (for any V).

Using the notations of Section 3, for $j \in \{0, \dots, L\}$ we define

$$S_{L,j}(E) := \sum_{k\neq j} \frac{a_k^{\mathbb{N}}}{\lambda_k - E} \quad\text{and}\quad \Gamma_{L,j}(E) := \sum_{k\neq j} \frac{1}{\lambda_k - E}\begin{pmatrix} |\varphi_k(L)|^2 & \overline{\varphi_k(0)}\varphi_k(L) \\ \varphi_k(0)\overline{\varphi_k(L)} & |\varphi_k(0)|^2 \end{pmatrix}. \tag{3-17}$$

We prove:

Theorem 3.3. *Pick $j \in \{0, \dots, L\}$ such that $-4 < \lambda_{j-1} + \lambda_j < \lambda_{j+1} + \lambda_j < 4$. There exists $C > 1$ (depending only on $(\lambda_{j-1} + \lambda_j) + 4$ and $4 - (\lambda_{j+1} + \lambda_j)$) such that, for any L, if $a_j \le d_j^2/C$, (2-4) and (2-8) have exactly one solution in the set*

$$R_j := \left\{ E \in \mathbb{C} \,\middle|\, \operatorname{Re} E \in \lambda_j + Ca_j[-1, 1],\ -Ca_j \le \operatorname{Im} E \le -\frac{a_j d_j^2}{C} \right\}. \tag{3-18}$$

Moreover, the solution to (2-4), say $z_j^{\mathbb{N}}$, satisfies

$$z_j^{\mathbb{N}} = \lambda_j + \frac{a_j^{\mathbb{N}}}{S_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)}} + O((a_j^{\mathbb{N}} d_j^{-1})^2) \tag{3-19}$$

and the solution to (2-8), say $z_j^{\mathbb{Z}}$, satisfies

$$z_j^{\mathbb{Z}} = \lambda_j + \left\langle \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix}, (\Gamma_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)})^{-1} \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix} \right\rangle + O((a_j^{\mathbb{Z}} d_j^{-1})^2). \tag{3-20}$$

Note that, if $a_j^{\mathbb{N}} d_j^{-2}$ is small, (3-19) gives the asymptotic of the width of the solution $z_j^{\mathbb{N}}$, namely,

$$\operatorname{Im} z_j^{\mathbb{N}} = \frac{a_j^{\mathbb{N}} \sin\theta(\lambda_j)}{[S_{L,j}(\lambda_j) + \cos\theta(\lambda_j)]^2 + \sin^2\theta(\lambda_j)}(1 + o(1)). \tag{3-21}$$

Recall that $\sin\theta(\lambda_j) < 0$ (see Theorem 2.2). For $H_L^{\mathbb{Z}}$, using the bounds (3-28) and (3-29), we see that the asymptotic of the imaginary part of the solution $z_j^{\mathbb{Z}}$ satisfies

$$-\frac{1}{C}a_j^{\mathbb{Z}} \le \operatorname{Im} z_j^{\mathbb{Z}} \le -Ca_j^{\mathbb{Z}}d_j^2. \tag{3-22}$$

This and (3-21) will be useful when $a_j^\bullet \ll d_j^2$, as will be the case for random potentials. The case when $a_j^\bullet$ and d_j are of the same order of magnitude requires more information. This is the case that we meet in the next section when dealing with periodic potentials.

The proof of Theorem 3.3 also yields the behavior of the functions $E \mapsto S_L(E) + e^{-i\theta(E)}$ and $E \mapsto \det(\Gamma_L(E) + e^{-i\theta(E)})$ near their zeros in R_j and, in particular, shows the following:

Proposition 3.4. *Fix $\delta > 0$. Under the assumptions of Theorem 3.3, there exists $c > 0$ such that, for $-4+\delta < \lambda_{j-1}+\lambda_j < \lambda_{j+1}+\lambda_j < 4-\delta$, one has*

$$\inf_{0<r<ca_j^{\mathbb{N}}d_j^{-1}} \min_{|E-z_j^{\mathbb{N}}|=r} \frac{|S_L(E)+e^{-i\theta(E)}|}{r} \ge c \quad \text{and} \quad \inf_{0<r<ca_j^{\mathbb{Z}}d_j^{-1}} \min_{|E-z_j^{\mathbb{Z}}|=r} \frac{|\det(\Gamma_L(E)+e^{-i\theta(E)})|}{r} \ge c.$$

Proposition 3.4 is a consequence of the analogues of (3-24) and (3-30) on the rectangles

$$\widetilde{R}_j = \tilde{z}_j + ca_j^\bullet d_j^{-1}[-1,1]\times[-1,1]$$

for $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$ and c sufficiently small.

Proof of Theorem 3.3. Let us start with (2-4). To prove the statement in (2-4), in R_j we compare the function $E \mapsto S_L(E) + e^{-i\theta(E)}$ to the function

$$E \mapsto \tilde{S}_{L,j}(E) = \frac{a_j^{\mathbb{N}}}{\lambda_j - E} + S_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)}.$$

Clearly, in $\mathbb{C}$, the equation $\tilde{S}_{L,j}(E) = 0$ admits a unique solution, given by

$$\tilde{z}_j = \lambda_j + \frac{a_j^{\mathbb{N}}}{S_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)}}.$$

For $E \in \partial R_j$, the boundary of R_j, one has

$$\begin{gathered} |\tilde{S}_{L,j}(E)| \ge \frac{1}{2C} \quad \text{and} \quad \left|\frac{a_j^{\mathbb{N}}}{\lambda_j - E}\right| \ge \frac{1}{2C}, \\ |e^{-i\theta(E)} - e^{-i\theta(\lambda_j)}| \le Ca_j^{\mathbb{N}} \quad \text{and} \quad |S_{L,j}(E) - S_{L,j}(\lambda_j)| \le Ca_j^{\mathbb{N}}d_j^{-2}. \end{gathered} \tag{3-23}$$

Hence, as $d_j \le 1$, one gets

$$\max_{E\in\partial R_j} \frac{|\tilde{S}_{L,j}(E) - S_L(E) - e^{-i\theta(E)}|}{|\tilde{S}_{L,j}(E)|} \le 4Ca_j^{\mathbb{N}}d_j^{-2}.$$

Thus, by Rouché's theorem, (2-4) has a unique solution in R_j.

To obtain the asymptotics of the solution, it suffices to use Rouché’s theorem again with the functions $\tilde{S}_{L,j}$ and $S_L(E)+e^{-i\theta(E)}$ on the smaller rectangle $\widetilde{R}_j=\tilde{z}_j+K(a_j^{\mathbb{N}}d_j^{-1})^2[-1,1]\times[-1,1]$. One then estimates

$$\max_{E\in\partial\widetilde{R}_j}\frac{|\tilde{S}_{L,j}(E)-S_L(E)-e^{-i\theta(E)}|}{|\tilde{S}_{L,j}(E)|}\le 4CK^{-1}. \tag{3-24}$$

Thus, for K sufficiently large, this completes the proof of the statements on the solutions to (2-4) contained in Theorem 3.3.

Let us turn to (2-8). On R_j, we now compare $\Gamma_L(E)+e^{-i\theta(E)}$ to the matrix-valued function

$$E\mapsto\widetilde{\Gamma}_{L,j}(E):=\frac{1}{\lambda_j-E}\begin{pmatrix}|\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2\end{pmatrix}+\Gamma_{L,j}(\lambda_j)+e^{-i\theta(\lambda_j)}.$$

The matrix

$$\begin{pmatrix}|\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2\end{pmatrix}$$

has rank 1 and can be diagonalized as

$$\begin{pmatrix}|\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2\end{pmatrix}=P_j\begin{pmatrix}a_j^{\mathbb{Z}} & 0\\ 0 & 0\end{pmatrix}P_j^*,$$

where $a_j^{\mathbb{Z}}$ is given by (2-13) and

$$P_j=\frac{1}{\sqrt{a_j^{\mathbb{Z}}}}\begin{pmatrix}\varphi_j(L) & -\overline{\varphi_j(0)}\\ \varphi_j(0) & \overline{\varphi_j(L)}\end{pmatrix}.$$

Thus, $\widetilde{\Gamma}_{L,j}(E)$ is unitarily equivalent to

$$M:=\frac{1}{\lambda_j-E}\begin{pmatrix}a_j^{\mathbb{Z}} & 0\\ 0 & 0\end{pmatrix}+P_j^*\Gamma_{L,j}(\lambda_j)P_j+e^{-i\theta(\lambda_j)}. \tag{3-25}$$

As $P_j^*\Gamma_{L,j}(\lambda_j)P_j$ is real and the imaginary part of $e^{-i\theta(\lambda_j)}$ does not vanish, $M_0:=P_j^*\Gamma_{L,j}(\lambda_j)P_j+e^{-i\theta(\lambda_j)}$ is invertible. By rank-1 perturbation theory (see, e.g., [Simon 2005]), we know that M is invertible if and only if $a_j^{\mathbb{Z}}[M_0^{-1}]_{11}+\lambda_j\neq E$ (where $[M]_{11}$ is the upper right coefficient of the 2×2 matrix M). In this case, one has

$$M^{-1}=M_0^{-1}-\frac{a_j^{\mathbb{Z}}}{a_j^{\mathbb{Z}}[M_0^{-1}]_{11}+\lambda_j-E}M_0^{-1}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}M_0^{-1}. \tag{3-26}$$

Hence, 0 is an eigenvalue of M if and only if

$$\begin{aligned}E&=\lambda_j+a_j^{\mathbb{Z}}[(P_j^*\Gamma_{L,j}(\lambda_j)P_j+e^{-i\theta(\lambda_j)})^{-1}]_{11}\\ &=\lambda_j+\left\langle\begin{pmatrix}\overline{\varphi_j(L)}\\ \overline{\varphi_j(0)}\end{pmatrix},(\Gamma_{L,j}(\lambda_j)+e^{-i\theta(\lambda_j)})^{-1}\begin{pmatrix}\overline{\varphi_j(L)}\\ \overline{\varphi_j(0)}\end{pmatrix}\right\rangle.\end{aligned} \tag{3-27}$$

Note that, as $\Gamma_{L,j}(\lambda_j)$ is real symmetric and $\|\Gamma_{L,j}(\lambda_j)\| \le C d_j^{-1}$, one has

$$\left|\left\langle \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix}, (\Gamma_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)})^{-1} \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix} \right\rangle\right| \le \frac{a_j^{\mathbb{Z}}}{|\sin\theta(\lambda_j)|} \tag{3-28}$$

and

$$\operatorname{Im}\left(\left\langle \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix}, (\Gamma_{L,j}(\lambda_j) + e^{-i\theta(\lambda_j)})^{-1} \begin{pmatrix} \overline{\varphi_j(L)} \\ \varphi_j(0) \end{pmatrix} \right\rangle\right) \le \frac{a_j^{\mathbb{Z}} d_j^2 \sin\theta(\lambda_j)}{1 + d_j^2}. \tag{3-29}$$

Using (3-25), (3-26), (3-28) and (3-29),we see that, for $E \in \partial R_j$, the boundary of R_j, $\widetilde{\Gamma}_{L,j}(E)$ is invertible and that one has

$$\|[\widetilde{\Gamma}_{L,j}(E)]^{-1}\| \le 2C \quad \text{and} \quad \|\Gamma_{L,j}(E) - \Gamma_{L,j}(\lambda_j)\| \le C a_j^{\mathbb{Z}} d_j^{-2}.$$

Hence, as $d_j \le 1$, taking (3-23) into account, one gets

$$\max_{E \in \partial R_j} \|1 - [\widetilde{\Gamma}_{L,j}(E)]^{-1}(\Gamma_L(E) + e^{-i\theta(E)})\| \le 4C^2 a_j^{\mathbb{Z}} d_j^{-2}.$$

In the same way, one proves

$$\max_{E \in \partial \widetilde{R}_j} \|1 - [\widetilde{\Gamma}_{L,j}(E)]^{-1}(\Gamma_L(E) + e^{-i\theta(E)})\| \lesssim K^{-1}, \tag{3-30}$$

where we recall that $\widetilde{R}_j = \tilde{z}_j + K(a_j^{\mathbb{N}} d_j^{-1})^2[-1, 1] \times [-1, 1]$.

Thus, we can apply Rouché's theorem to compare the following two functions on ∂R_j and $\partial \widetilde{R}_j$ (for K sufficiently large):

$$\det(\widetilde{\Gamma}_{L,j}(E)) \quad \text{and} \quad \det(\Gamma_L(E) + e^{-i\theta(E)}),$$

as

$$\frac{|\det(\widetilde{\Gamma}_{L,j}(E)) - \det(\Gamma_L(E) + e^{-i\theta(E)})|}{|\det(\widetilde{\Gamma}_{L,j}(E))|} = \left|1 - \det\left(1 - \left[1 - [\widetilde{\Gamma}_{L,j}(E)]^{-1}(\Gamma_L(E) + e^{-i\theta(E)})\right]\right)\right|.$$

We then conclude as in the case of (2-4). This completes the proof of Theorem 3.3. □

Combining Theorems 3.3, 3.1 and 3.2, we get a pretty clear picture of the resonances near the Dirichlet eigenvalues in $(-2, 2)$ as long as the associated $a_j^{\bullet}$ and d_j behave correctly. As said, this and the knowledge of the spectral statistics for random operators will enable us to prove the results described in Section 1C. For the periodic case, Theorems 3.1, 3.2 and 3.3 will prove not to be sufficient. As we shall see, in this case, $a_j^{\bullet}$ and d_j are of the same order of magnitude. Thus, neighboring Dirichlet eigenvalues have a sizable effect on the location of resonances. Therefore, in the next section, we compute the Dirichlet spectral data for the truncated periodic potential.

4. The Dirichlet spectral data for periodic potentials

As we did not find any suitable reference for this material, we first derive a suitable description of the spectral data (i.e., the $(a_j^{\bullet})_j$ and $(\lambda_j)_j$) for the Dirichlet restriction of a periodic operator to the interval $[\![0, L]\!]$ when L becomes large.

Consider a potential $V : \mathbb{N} \to \mathbb{R}$ such that, for some $p \geq 1$, one has $V_k = V_{k+p}$ for all $k \geq 0$. We assume p to be minimal, i.e., to be the period of V. In our first result, we describe the spectrum of $H^{\mathbb{Z}} = -\Delta + V$ on $\ell^2(\mathbb{Z})$ and $H^{\mathbb{N}} = -\Delta + V$ on $\ell^2(\mathbb{N})$ (with Dirichlet boundary conditions at 0). In the second result we turn to H_L, the Dirichlet restriction $H^{\mathbb{N}}$ to $[\![0, L]\!]$ and describe its spectral data, i.e., its eigenvalues and eigenfunctions.

We recall:

Theorem 4.1. *The spectrum of $H^{\mathbb{Z}}$, say $\Sigma_{\mathbb{Z}}$, is a union of at most p disjoint intervals that all consist in purely absolutely continuous spectrum.*

The spectrum of $H^{\mathbb{N}}$ is the union of $\Sigma_{\mathbb{Z}}$ and at most finitely many simple eigenvalues outside $\Sigma_{\mathbb{Z}}$, say $(v_j)_{0 \leq j \leq n}$. $\Sigma_{\mathbb{Z}}$ consists of purely absolutely continuous spectrum and the eigenfunctions associated to $(v_j)_{0 \leq j \leq n}$, say $(\psi_j)_{0 \leq j \leq n}$, are exponentially decaying at infinity.

Except for the exponential decay of the eigenfunctions, the proof of the statement for the periodic operator on $\mathbb{Z}$ and $\mathbb{N}$ is classical and can, e.g., be found in a more general setting in [Teschl 2000, Chapters 2, 3 and 7] (see also [van Moerbeke 1976; Reed and Simon 1980]). The exponential decay is an immediate consequence of Floquet theory for the periodic Hamiltonian on $\mathbb{Z}$ and the fact that the eigenvalues lie in gaps of $\Sigma_{\mathbb{Z}}$.

For $H^{\mathbb{Z}}$, one can define its Bloch quasimomentum (see the beginning of Section 4A for details), which we denote by θ_p; it is continuous and strictly increasing on $\Sigma_{\mathbb{Z}}$ and real analytic on $\Sigma_{\mathbb{Z}}^{\circ}$, the interior of $\Sigma_{\mathbb{Z}}$. Decompose $\Sigma_{\mathbb{Z}}$ into its connected components, i.e., $\Sigma_{\mathbb{Z}} = \bigcup_{r=1}^{q} B_r$, where $q \leq p$. Let c_q be the number of closed gaps contained in q. Then θ_p is continuous and strictly increasing on B_r and real analytic on B_r°, the interior of the r-th band. Moreover, on this set, its derivative can be expressed in terms of the density of states, defined in (1-2) as

$$n(\lambda) = \frac{1}{\pi}\theta_p'(\lambda). \tag{4-1}$$

We first describe the eigenvalues of H_L.

Theorem 4.2. *One has*:

(1) *For any $k \in \{0, \dots, p-1\}$, there exists $h_k : \Sigma_{\mathbb{Z}} \to \mathbb{R}$, a continuous function that is real analytic in a neighborhood of $\Sigma_{\mathbb{Z}}^{\circ}$ such that, for L sufficiently large with $L \equiv k \bmod p$,*

  (a) *for $1 \leq r \leq q$, the function h_k maps B_r into $(-(c_r+1)\pi, (c_r+1)\pi)$;*

  (b) *the function*

$$\theta_{p,L} := \theta_p - \frac{h_k}{L-k} \tag{4-2}$$

  is continuous and strictly monotonous on each B_r $(1 \leq r \leq q)$;

  (c) *for $1 \leq r \leq q$, the eigenvalues of H_L in B_r, the r-th band of $\Sigma_{\mathbb{Z}}$, say $(\lambda_j^r)_j$, are the solutions (in $\Sigma_{\mathbb{Z}}$) to the quantization conditions*

$$\theta_{p,L}(\lambda_j^r) = \frac{j\pi}{L-k}, \quad j \in \mathbb{Z}. \tag{4-3}$$

(2) *There exists* $c > 0$ *such that, if* λ *is an eigenvalue of* H_L *outside* $\Sigma_{\mathbb{Z}}$*, then for* $L = Np + k$ *sufficiently large there exists* $\lambda_\infty \in \Sigma_0^+ \cup \Sigma_k^- \setminus \Sigma_{\mathbb{Z}}$ *such that one has* $|\lambda - \lambda_\infty| \le e^{-cL}$.

Recall that Σ_0^+ and Σ_k^- are the spectra of H_0^+ and H_k^-, respectively, defined in Section 1B2.

In Theorem 4.2, when solving (4-3), one has to do it for each band B_r and, for each band and each j such that $j\pi/(L-k) \in \theta_{p,L}(B_r)$, (4-3) admits a unique solution. But, it may happen that one has two solutions to (4-3) for a given j belonging to neighboring bands. In the sequel, to simplify the notations, we will not distinguish between the different bands, i.e., we will write eigenvalues $(\lambda_j)_j$ not referring to the band they belong to.

Let us now describe the associated eigenfunctions.

Theorem 4.3. *Recall that* $(\lambda_j)_j$ *are the eigenvalues of* H_L *in* $\Sigma_{\mathbb{Z}}$ *(enumerated as in Theorem 4.2).*

(1) *There exist* $p+2$ *positive functions, say* f_0^+, $(f_k^-)_{0\le k\le p-1}$ *and* $\tilde f$, *that are real analytic in a neighborhood of* $\Sigma_{\mathbb{Z}}^\circ$ *such that there exists* $\sigma_r \in \{+1, -1\}$ *such that, for* $L = Np + k$ *sufficiently large and* λ_j *in* B_r°*, the interior of r-th band of* $\Sigma_{\mathbb{Z}}$*, one has*

$$|\varphi_l(L)|^2 = \frac{f_k^-(\lambda_j)}{L-k}\left(1 + \frac{\tilde f(\lambda_j)}{L-k}\right)^{-1}, \quad |\varphi_l(0)|^2 = \frac{f_0^+(\lambda_j)}{f_k^-(\lambda_j)}|\varphi_l(L)|^2,$$
$$\textit{and} \quad \varphi_l(L)\overline{\varphi_l(0)} = \sigma_r e^{i\pi l}|\varphi_l(L)|\,|\varphi_l(0)| = \sigma_r e^{i(L-k)\theta_p(\lambda_j) - h_k(\lambda_j)}|\varphi_l(L)|\,|\varphi_l(0)|. \tag{4-4}$$

(2) *Let* λ *be an eigenvalue of* H_L *outside* $\Sigma_{\mathbb{Z}}$ *(see Theorem 4.2(2)). If* φ *is a normalized eigenfunction associated to* λ *and* H_L*, one has one of the following alternatives for* L *large*:

(a) *If* $\lambda_\infty \in \Sigma_0^+ \setminus \Sigma_k^-$*, one has*

$$|\varphi(L)| \asymp e^{-cL} \quad \textit{and} \quad |\varphi(0)| \asymp 1. \tag{4-5}$$

(b) *If* $\lambda_\infty \in \Sigma_k^- \setminus \Sigma_0^+$*, one has*

$$|\varphi(L)| \asymp 1 \quad \textit{and} \quad |\varphi(0)| \asymp e^{-cL}. \tag{4-6}$$

(c) *If* $\lambda_\infty \in \Sigma_k^- \cap \Sigma_0^+$*, one has*

$$|\varphi(L)| \asymp 1 \quad \textit{and} \quad |\varphi(0)| \asymp 1. \tag{4-7}$$

For later use, let us define $\theta_{p,L}$, $f_{0,L}$ and $f_{k,L}$ by

$$f_{k,L}(\lambda) = f_k^-(\lambda)\left(1 + \frac{\tilde f(\lambda)}{L-k}\right)^{-1} \quad \text{and} \quad f_{0,L}(\lambda) = f_0^+(\lambda)\left(1 + \frac{\tilde f(\lambda)}{L-k}\right)^{-1}, \tag{4-8}$$

where θ_p, h_k, f_0, f_k and $\tilde f$ are as defined in Theorem 4.2.

As a consequence of Theorem 4.2, we obtain:

Corollary 4.4. *For $\lambda\in\Sigma_{\mathbb{Z}}^{\circ}$, for $L\equiv k \bmod p$ sufficiently large, one has*

$$\frac{dN_k^-}{d\lambda}(\lambda)=n_k^-(\lambda)=f_k^-(\lambda)n(\lambda)=\frac{1}{\pi}f_k^-(\lambda)\theta_p'(\lambda)=\frac{1}{\pi}f_{k,L}(\lambda)\theta_{p,L}'(\lambda), \tag{4-9}$$

$$\frac{dN_0^+}{d\lambda}(\lambda)=n_0^+(\lambda)=f_0^+(\lambda)n(\lambda)=\frac{1}{\pi}f_0^+(\lambda)\theta_p'(\lambda)=\frac{1}{\pi}f_{0,L}(\lambda)\theta_{p,L}'(\lambda). \tag{4-10}$$

Here, θ_P, f_0^+ and f_k^- are the functions defined in Theorem 4.2.

Proof of Corollary 4.4. To prove the first equalities in (4-9) and (4-10), it suffices to prove that, for any $\chi\in\mathscr{C}_0^\infty(\Sigma_{\mathbb{Z}}^{\circ})$,

$$\langle\delta_0,\chi(H_k^-)\delta_0\rangle=\int_{\mathbb{R}}\chi(\lambda)\,dN_k^-(\lambda)=\frac{1}{\pi}\int_{\mathbb{R}}\chi(\theta_p^{-1}(k))f_k^-(\theta_p^{-1}(k))\,dk=\frac{1}{\pi}\int_{\mathbb{R}}\chi(\lambda)f_k^-(\lambda)\theta_p'(\lambda)\,d\lambda, \tag{4-11}$$

$$\langle\delta_0,\chi(H_0^+)\delta_0\rangle=\int_{\mathbb{R}}\chi(\lambda)\,dN_0^+(\lambda)=\frac{1}{\pi}\int_{\mathbb{R}}\chi(\theta_p^{-1}(k))f_0^+(\theta_p^{-1}(k))\,dk=\frac{1}{\pi}\int_{\mathbb{R}}\chi(\lambda)f_0^+(\lambda)\theta_p'(\lambda)\,d\lambda, \tag{4-12}$$

the full statement then following by standard density argument. The operator H_L converges to H_0^+ in the norm resolvent sense. Thus, we know that $\langle\delta_0,\chi(H_0^+)\delta_0\rangle=\lim_{L\to+\infty}\langle\delta_0,\chi(H_L)\delta_0\rangle$. Now, by Theorem 4.2, as χ is supported in $\Sigma_{\mathbb{Z}}^{\circ}$, using the Poisson formula one computes

$$\begin{aligned}\langle\delta_0,\chi(H_L)\delta_0\rangle&=\sum_j\chi(\lambda_j)|\,|\varphi_j(0)|^2=\frac{1}{L-k}\sum_l\chi\left(\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right)\right)f_{0,L}\left(\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right)\right)\\&=\frac{1}{L-k}\sum_{j\in\mathbb{Z}}\int_{\mathbb{R}}e^{-i2\pi j\lambda}\chi\left(\theta_{p,L}^{-1}\left(\frac{\pi\lambda}{L-k}\right)\right)f_{0,L}\left(\theta_{p,L}^{-1}\left(\frac{\pi\lambda}{L-k}\right)\right)d\lambda\\&=\frac{1}{\pi}\sum_{j\in\mathbb{Z}}\int_{\mathbb{R}}e^{-i2(L-k)j\theta_{p,L}(\lambda)}\chi(\lambda)f_{0,L}(\lambda)\theta_{p,L}'(\lambda)\,d\lambda.\end{aligned}$$

Thus, using the nonstationary phase, i.e., integrating by parts, one gets, for any $N\geq 2$,

$$\begin{aligned}\left|\langle\delta_0,\chi(H_L)\delta_0\rangle-\frac{1}{\pi}\int_{\mathbb{R}}\chi(\lambda)f_{0,L}(\lambda)\theta_{p,L}'(\lambda)\,d\lambda\right|&\leq\sum_{j\geq1}C_{N,K}\|\chi\|_{\mathscr{C}^N}(|j|(L-k))^{-N}\\&\leq C_{N,K}\|\chi\|_{\mathscr{C}^N}((L-k))^{-N}.\end{aligned} \tag{4-13}$$

Here we have used the analyticity of the functions $\theta_{p,L}$ and $f_{0,L}$.

To deal with H_k^-, we recall the operator $\widetilde{H}_L$ (which is unitarily equivalent to H_L) defined in Remark 1.6. One has $\langle\delta_L,H_L\delta_L\rangle=\langle\delta_0,\chi(\widetilde{H}_L)\delta_0\rangle$; thus, as H_k^- is the strong resolvent sense limit of $\widetilde{H}_L$, one gets $\langle\delta_0,\chi(H_k^-)\delta_0\rangle=\lim_{L\to+\infty}\langle\delta_L,\chi(H_L)\delta_L\rangle$.

Then (4-11) and (4-12) — and, thus, the first equalities in (4-9) and (4-10) — follow, as $\theta_{p,L}'$, $f_{0,L}$ and $f_{k,L}$ converge (locally uniformly on $\Sigma_{\mathbb{Z}}^{\circ}$) to θ_p', f_0^+ and f_k^-, respectively (see (4-8) and Theorem 4.2).

Let us now prove the second equalities in (4-9) and (4-10). To this end, we use an *almost analytic extension* (see [Mather 1971]) of χ, say $\tilde{\chi}$, that is, a function $\tilde{\chi}:\mathbb{C}\to\mathbb{C}$ satisfying

(1) $\tilde{\chi}(z) = \chi(z)$ for $z \in \mathbb{R}$,

(2) $\operatorname{supp}(\tilde{\chi}) \subset \{z \in \mathbb{C} \mid |\operatorname{Im}(z)| < 1\}$,

(3) $\tilde{\chi} \in \mathscr{S}(\{z \in \mathbb{C} \mid |\operatorname{Im}(z)| < 1\})$,

(4) the family of functions $x \mapsto \partial\tilde{\chi}(x+iy)/\partial\bar{z} \cdot |y|^{-n}$ (for $0 < |y| < 1$) is bounded in $\mathscr{S}(\mathbb{R})$ for any $n \in \mathbb{N}$.

Moreover, $\tilde{\chi}$ can be chosen so that one has the following estimates: for $n \geq 0$, $\alpha \geq 0$, $\beta \geq 0$, there exists $C_{n,\alpha,\beta} > 0$ such that

$$\sup_{0<|y|\leq 1} \sup_{x\in\mathbb{R}} \left| x^\alpha \frac{\partial^\beta}{\partial x^\beta} \left(|y|^{-n} \frac{\partial\tilde{\chi}}{\partial\bar{z}}(x+iy) \right) \right| \leq C_{n,\alpha,\beta} \sup_{\beta'\leq n+\beta+2\alpha'\leq\alpha} \sup_{x\in\mathbb{R}} \left| x^{\alpha'} \frac{\partial^{\beta'}\chi}{\partial x^{\beta'}}(x) \right|. \tag{4-14}$$

By the definition of χ, the right-hand side of (4-14) is bounded uniformly in E complex.

Let $\chi \in \mathscr{C}_0^\infty(\mathbb{R})$ and $\tilde{\chi}$ be an almost analytic extension of $\chi(x)$. Then, by [Helffer and Sjöstrand 1990; Klopp 1995], we know that, for any n and $\omega \in \Omega$,

$$\chi(H_\bullet) = \frac{i}{2\pi} \int_{\mathbb{C}} \frac{\partial\tilde{\chi}}{\partial\bar{z}}(z)(z - H_\bullet)^{-1}\, dz \wedge d\bar{z}, \tag{4-15}$$

where $H_\bullet$ equals H_L, $\widetilde{H}_L$, H_0^+ or H_k^-.

Using the geometric resolvent equation (see, e.g., [Kirsch 2008, Theorem 5.20]) and the Combes–Thomas estimate (see, e.g., [Kirsch 2008, Theorem 11.2]), we know that for some $C > 0$, for $\operatorname{Im} z \neq 0$,

$$\left|\langle \delta_0, [(\widetilde{H}_L - z)^{-1} - (H_k^- - z)^{-1}]\delta_0\rangle\right| + \left|\langle \delta_0, [(H_L - z)^{-1} - (H_0^+ - z)^{-1}]\delta_0\rangle\right| \leq \frac{C}{|\operatorname{Im} z|} e^{-L|\operatorname{Im} z|/C}. \tag{4-16}$$

Plugging (4-16) into (4-15) and using (4-14), we get

$$\left| \sum_{j=0}^{L} \chi(\lambda_j)|\varphi_j(0)|^2 - \int_{\mathbb{R}} \chi(\lambda)\, dN_0^+(\lambda) \right| \leq \widetilde{C}_N \int_{|y|\leq 1} |y|^{N-1} e^{-L|y|/C}\, dy \leq C_N L^{-N}.$$

Thus, by (4-12) and (4-13), we obtain that, for $\chi \in \mathscr{C}_0^\infty(\Sigma_{\mathbb{Z}}^\circ)$ and any $N \geq 0$, there exists $C_N > 0$ such that

$$\begin{aligned} &\left| \int_{\mathbb{R}} \chi(\lambda)[f_{0,L}(\lambda)\theta'_{p,L}(\lambda) - f_0^+(\lambda)\theta'_p(\lambda)]\, d\lambda \right| \\ &\qquad = \left| \int_{\mathbb{R}} \chi(\lambda) f_{0,L}(\lambda)\theta'_{p,L}(\lambda)\, d\lambda - \int_{\mathbb{R}} \chi(\lambda)\, dN_0^+(\lambda) \right| \leq C_N L^{-N}. \end{aligned} \tag{4-17}$$

Now, by (4-3) and (4-8), the function $f_{0,L}\theta'_{p,L} - f_0^+\theta'_p$ admits an expansion in inverse powers of L that converges uniformly on compact subsets of $\Sigma_{\mathbb{Z}}^\circ$, namely,

$$f_{0,L}\theta'_{p,L} - f_0^+\theta'_p = \sum_{k\geq 1} L^{-k}\alpha_k.$$

Thus, (4-17) immediately yields that, for any $k \geq 1$, one has $\alpha_k \equiv 0$ on $\Sigma_{\mathbb{Z}}^\circ$. Hence, $f_{0,L}\theta'_{p,L} \equiv f_0^+\theta'_p$ on $\Sigma_{\mathbb{Z}}^\circ$. This completes the proof of Corollary 4.4. □

4A. ***The proofs of Theorems 4.2 and 4.3.*** We will first describe some objects from the spectral theory of $H^{\mathbb{Z}}$, use them to describe the spectral theory of $H^{\mathbb{N}}$, prove Theorem 4.2 and finally prove Theorem 4.3.

4A1. *The spectral theory of* $H^{\mathbb{Z}}$. This material is classical (see, e.g., [van Moerbeke 1976; Teschl 2000]); we only recall it for the reader's convenience. For $0 \le j \le p-1$, define $\widetilde{T}_j = \widetilde{T}_j(E)$ to be a monodromy matrix for the periodic finite difference operator $H^{\mathbb{Z}}$, that is,

$$\widetilde{T}_j(E) = T_{j+p-1,j}(E) = T_{j+p-1}(E)\cdots T_j(E) =: \begin{pmatrix} a_p^j(E) & b_p^j(E) \\ a_{p-1}^j(E) & b_{p-1}^j(E) \end{pmatrix}, \tag{4-18}$$

where

$$T_j(E) = \begin{pmatrix} E - V_j & -1 \\ 1 & 0 \end{pmatrix}. \tag{4-19}$$

The coefficients of $\widetilde{T}_j(E)$ are monic polynomials in the energy E; $a_p^j(E)$ has degree p and $b_p^j(E)$ has degree $p-1$. Clearly, $\det \widetilde{T}_j(E) = 1$. As $j \mapsto V_j$ is p-periodic, so is $j \mapsto \widetilde{T}_j(E)$. Moreover, for $j' < j$, one has

$$\widetilde{T}_j(E) T_{j,j}(E) = T_{j+p-1,j'+p-1}(E)\widetilde{T}_{j'}(E) = T_{j,j'}(E)\widetilde{T}_{j'}(E). \tag{4-20}$$

Thus, the discriminant $\underline{\Delta}(E) := \operatorname{tr} \widetilde{T}_j(E) = a_p^j(E) + b_{p-1}^j(E)$ is a polynomial of degree p that is independent of j; so are $\rho(E)$ and $\rho^{-1}(E)$, the eigenvalues of $\widetilde{T}_j(E)$. One defines the Bloch quasimomentum $E \mapsto \theta_p(E)$ by

$$\underline{\Delta}(E) = \rho(E) + \rho^{-1}(E) = 2\cos(p\theta_p(E)). \tag{4-21}$$

Let us recall some basic properties of the discriminant Δ and the coefficients of $\widetilde{T}_j$, the proofs of which can be found in [van Moerbeke 1976]:

(1) If $\Delta'(E) = 0$ then $|\underline{\Delta}(E)| \ge 2$.

(2) The zeros of Δ' are simple.

(3) E is a zero of Δ' such that $|\underline{\Delta}(E)| = 2$ if and only if $\widetilde{T}_j(E) \in \{+\mathrm{Id}, -\mathrm{Id}\}$ (for any j).

(4) The polynomials b_p^j and a_{p-1}^j only vanish in the set $\{|\underline{\Delta}(E)| \ge 2\}$; they keep a fixed sign in each of the connected components of the set $\{|\underline{\Delta}(E)| < 2\}$.

Note that $\underline{\Delta}(E)$ is real when E is real. Thus, for E real, $|\underline{\Delta}(E)| \le 2$ implies that $\rho^{-1}(E) = \overline{\rho(E)}$ and $|\underline{\Delta}(E)| > 2$ implies that $\rho(E)$ is real. When $|\underline{\Delta}(E)| \le 2$ we will fix $\rho(E) := e^{ip\theta_p(E)}$ and when $|\underline{\Delta}(E)| > 2$ we will fix $\rho(E)$ so that $|\rho(E)| < 1$.

E belongs to the spectrum of $H^{\mathbb{Z}}$ (i.e., $-\Delta + V$ on $\ell^2(\mathbb{Z})$) if and only if $|\underline{\Delta}(E)| \le 2$ (see, e.g., [Teschl 2000]).

Properties (1)–(3) above imply that, for E_0 a zero of Δ' such that $\underline{\Delta}(E_0) = \pm 2$, θ_p is real analytic near E_0 and $\theta_p'(E_0) \ne 0$.

Definition 4.5. E_0 is said to be a closed gap if and only if $|\underline{\Delta}(E_0)| = 2$ and $\Delta'(E_0) = 0$ or, equivalently, if and only if $\widetilde{T}_0(E_0)$ is diagonal.

Consider $\partial\Sigma_{\mathbb{Z}}$. It is the set of energies that are solutions to $|\underline{\Delta}(E)|=2$ where $\widetilde{T}_0(E)$ is not diagonal; it is also the set of roots of $|\underline{\Delta}(E)|=2$ that are not closed gaps. From the upper half of the complex plane, one can continue $E\mapsto\theta_p(E)$ analytically to the universal cover of $\mathbb{C}\setminus\partial\Sigma_{\mathbb{Z}}$. Each of the points in $\partial\Sigma_{\mathbb{Z}}$ is a branch point of θ_p of square root type. Moreover, for $E\notin\partial\Sigma_{\mathbb{Z}}$, there exist two linearly independent solutions to the eigenvalue equation $(-\Delta+V-E)u=0$, say $\varphi_\pm(E)$, satisfying, for $n\in\mathbb{Z}$,

$$\varphi_\pm(n+p,E)=e^{\pm ip\theta_p(E)}\varphi_\pm(n,E). \tag{4-22}$$

4A2. *The spectral theory of $H^{\mathbb{N}}$.* Let us now turn to the spectrum of the operator on the half-lattice.

The operator H_0^+. For the operator $H_0^+=H^{\mathbb{N}}$ (that is, $-\Delta+V$ on $\ell^2(\mathbb{N})$ with Dirichlet boundary conditions at 0), E is in the spectrum if and only if

- either $|\underline{\Delta}(E)|\le 2$,
- or $|\underline{\Delta}(E)|>2$ and $[\widetilde{T}_0(E)]^n\binom{1}{0}$ stays bounded as $n\to+\infty$.

The second condition is equivalent to requiring that $[\widetilde{T}_j(E)]^nT_{j-1}(E)\cdots T_0(E)\binom{1}{0}$ stay bounded as $n\to+\infty$.

When $|\underline{\Delta}(E)|\neq 2$ and $a^0_{p-1}(E)\neq 0$, one can diagonalize $\widetilde{T}_0(E)$ in the following way

$$\begin{pmatrix} a^0_{p-1}(E) & \rho(E)-a^0_p(E)\\ -a^0_{p-1}(E) & a^0_p(E)-\rho^{-1}(E)\end{pmatrix}\widetilde{T}_0(E)=\begin{pmatrix}\rho(E) & 0\\ 0 & \rho^{-1}(E)\end{pmatrix}\begin{pmatrix} a^0_{p-1}(E) & \rho(E)-a^0_p(E)\\ -a^0_{p-1}(E) & a^0_p(E)-\rho^{-1}(E)\end{pmatrix}. \tag{4-23}$$

Thus, using

$$\begin{vmatrix}\rho(E)-a^0_p(E) & -b^0_p(E)\\ -a^0_{p-1}(E) & \rho(E)-b^0_{p-1}(E)\end{vmatrix}=\begin{vmatrix}\rho(E)-a^0_p(E) & -b^0_p(E)\\ -a^0_{p-1}(E) & a^0_p(E)-\rho^{-1}(E)\end{vmatrix}=0 \tag{4-24}$$

for $n\in\mathbb{Z}$, one computes

$$(\widetilde{T}_0(E))^n=\begin{pmatrix}\tilde t^{11}_{0,n}(E) & \tilde t^{12}_{0,n}(E)\\ \tilde t^{21}_{0,n}(E) & \tilde t^{22}_{0,n}(E)\end{pmatrix}, \tag{4-25}$$

where

$$\begin{aligned}
\tilde t^{11}_{0,n}(E)&:=\rho^n(E)\frac{a^0_p(E)-\rho^{-1}(E)}{\rho(E)-\rho^{-1}(E)}+\rho^{-n}(E)\frac{\rho(E)-a^0_p(E)}{\rho(E)-\rho^{-1}(E)},\\
\tilde t^{12}_{0,n}(E)&:=(\rho^{-n}(E)-\rho^n(E))\frac{b^0_p(E)}{\rho(E)-\rho^{-1}(E)},\\
\tilde t^{21}_{0,n}(E)&:=(\rho^{n}(E)-\rho^{-n}(E))\frac{a^0_{p-1}(E)}{\rho(E)-\rho^{-1}(E)},\\
\tilde t^{22}_{0,n}(E)&:=\rho^{-n}(E)\frac{a^0_p(E)-\rho^{-1}(E)}{\rho(E)-\rho^{-1}(E)}+\rho^n(E)\frac{\rho(E)-a^0_p(E)}{\rho(E)-\rho^{-1}(E)}.
\end{aligned} \tag{4-26}$$

Clearly, the formulas (4-23), (4-25) and (4-26) stay valid even if $a^0_{p-1}(E)=0$. They also stay valid if $|\underline{\Delta}(E)|=2$ and $\Delta'(E)=0$. Indeed, by points (1)–(3) in Section 4A1, the functions $\rho-\rho^{-1}$, $a^0_p-\rho^{-1}$, $-\rho-a^0_p$, b^0_p and a^0_{p-1} are analytic near and have simple zeros at such points.

We have thus proved:

Lemma 4.6. *For $E \notin \partial\Sigma_{\mathbb{Z}}$, $(\widetilde{T}_0(E))^n$ has the form* (4-25)–(4-26)

Simple computations then show that E is in the spectrum of H_0^+ if and only if one of the following conditions is satisfied:

(1) $|\underline{\Delta}(E)| \le 2$: moreover, the set $\{E \in \mathbb{R} \mid |\underline{\Delta}(E)| \le 2\}$ is contained in the absolutely continuous spectrum of H_0^+.

(2) $|\underline{\Delta}(E)| > 2$ and

$$a_{p-1}^0(E) = 0 \quad \text{and} \quad |a_p^0(E)| < 1. \tag{4-27}$$

Thus, on $\Sigma_{\mathbb{Z}}$, the spectrum of H_0^+ is purely absolutely continuous; it does not contain any embedded eigenvalues.

Note that, in case (2), $[\widetilde{T}_0(E)]^n \binom{1}{0}$ actually decays exponentially fast. In this case, E is an eigenvalue associated to the (nonnormalized) eigenfunction $(u_l)_{l\in\mathbb{N}}$, where, for $n \ge 0$ and $j \in \{0, \dots, p-1\}$,

$$u_{np+j}(E) = \left\langle T_{j-1}(E)\cdots T_0(E)\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\rangle \cdot [a_p^0(E)]^n = a_j(E)[a_p^0(E)]^n, \tag{4-28}$$

writing

$$T_{j-1}(E)\cdots T_0(E) =: \begin{pmatrix} a_j(E) & b_j(E) \\ a_{j-1}(E) b_{j-1}(E) & \end{pmatrix}. \tag{4-29}$$

It is well known that, for any j, the zeros of a_j and b_j are simple (see, e.g., [Teschl 2000, Section 4]), and the roots of a_{j+1} (resp. b_{j+1}) interlace those of a_j (resp. b_j). Let E' be an eigenvalue of H_0^+. Differentiating (4-24) at the energy E', we compute

$$b_p^0(E')\frac{da_{p-1}^0}{dE}(E') + (\rho(E') - \rho^{-1}(E'))\frac{d(\rho - a_p^0)}{dE}(E') = 0. \tag{4-30}$$

The eigenvalues of the operator H_k^-. Let us now turn to H_k^-. Recalling (4-29) and using the representation (4-25), we obtain that the eigenvalues of H_k^- outside $\Sigma_{\mathbb{Z}}$ satisfy

$$\begin{pmatrix} \rho(E) - a_p^0(E) & -a_{p-1}^0(E) \\ -b_p^0(E) & a_p^0(E) - \rho^{-1}(E) \end{pmatrix}\begin{pmatrix} a_{k+1}(E) \\ b_{k+1}(E) \end{pmatrix} = 0. \tag{4-31}$$

As for H_0^+, the eigenfunction associated to E and H_k^- decays exponentially fast. Indeed, the eigenvalues of H_k^- in the region $|\underline{\Delta}(E)| > 2$ can be analyzed in the same way as we analyzed those of H_0^+, i.e., they are the energies such that $[\widetilde{T}_k(E)]^{-n}\binom{0}{1}$ stays bounded; this yields the quantization conditions $b_p^k(E) = 0$ and $|b_{p-1}^k(E)| < 1$. In this case, E is an eigenvalue associated to the (nonnormalized) eigenfunction $(u_l)_{-l\in\mathbb{N}}$, where, for $n \ge 0$ and $k \in \{0, \dots, p-1\}$,

$$u_{-np-k}(E) = b_k(E)[b_{p-1}^k(E)]^{-n}. \tag{4-32}$$

Common eigenvalues to H_0^+ and H_k^-. Assume now that E' is simultaneously an eigenvalue of H_k^- and H_0^+. In this case, one has $a_{p-1}^0(E')=0$, $|a_p^0(E')|<1$ and $b_p^0(E')b_{k+1}(E')=a_{k+1}(E')(\rho^{-1}(E')-\rho(E'))$. So (4-31) (see also (4-30)) becomes

$$\begin{pmatrix} d(\rho-a_p^0)(E')/dE & -da_{p-1}^0(E')/dE \\ -b_p^0(E) & a_p^0(E')-\rho^{-1}(E') \end{pmatrix}\begin{pmatrix} a_{k+1}(E') \\ b_{k+1}(E') \end{pmatrix}=0. \tag{4-33}$$

Hence, the analytic function $E\mapsto a_{k+1}(E)(a_p^0(E)-\rho(E))-b_{k+1}(E)a_{p-1}^0(E)$ has a root of order at least 2 at E'. It also implies that $a_{k+1}(E')\neq 0$. Indeed, if $a_{k+1}(E')=0$, (4-33) implies $b_{k+1}(E')=0$ as $da_{p-1}^0(E')/dE\neq 0$.

Conversely, if $E'\in\sigma(H_0^+)$ is such that $|\underline{\Delta}(E')|>2$ and $E\mapsto a_{k+1}(E)(a_p^0(E)-\rho(E))-b_{k+1}(E)a_{p-1}^0(E)$ has a root of order at least 2 at E', then (4-33) holds and E' is an eigenvalue of H_k^-.

We have thus proved:

Lemma 4.7. *$E_0\in\sigma(H_0^+)\cap\sigma(H_k^-)\setminus\mathbb{Z}$ if and only if $|\underline{\Delta}(E_0)|>2$ and E_0 is a double root of $E\mapsto a_{k+1}(E)(a_p^0(E)-\rho(E))-b_{k+1}(E)a_{p-1}^0(E)$.*

4A3. *The Dirichlet eigenvalues for a periodic potential: the proof of Theorem 4.2.* Let us now turn to the study of the eigenvalues and eigenvectors of H_L, i.e., to the proof of Theorem 4.2. We first prove the statements for the eigenvalues and then, in the next section, turn to the eigenvectors.

Recall that $L\equiv k \bmod p$; we write $L=Np+k$. By definition, E is an eigenvalue of $-\Delta+V$ on $[\![0,L]\!]$ with Dirichlet boundary conditions if and only if

$$\begin{aligned} 0&=\det\left(T_{L+1}(E)T_L(E)T_{L-1}(E)\cdots T_0(E)\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)\\ &=\det\left(T_k(E)\cdots T_0(E)\cdot[\widetilde{T}_0(E)]^N\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right), \end{aligned} \tag{4-34}$$

where $\widetilde{T}_k(E)$ is the monodromy matrix defined above.

We use the notations of sections 4A2 and 4A1. Let us first show Theorem 4.2(1), namely:

Lemma 4.8. *For L large, one has*

$$\partial\Sigma_{\mathbb{Z}}\cap\sigma(H_L)=\{E_0\mid a_{k+1}(E_0)=a_{p-1}^0(E_0)=0 \text{ and } b_p^0(E_0)\neq 0\}.$$

Proof. For $E_0\in\partial\Sigma_{\mathbb{Z}}$, we know that $|\underline{\Delta}(E_0)|=2$ and $\widetilde{T}_0(E_0)$ is not diagonal. Assume $\underline{\Delta}(E_0)=2$ (the case $\underline{\Delta}(E_0)=-2$ is dealt with in the same way); hence, $\widetilde{T}_0(E_0)$ has a Jordan normal form, i.e., there exists a 2×2 invertible matrix P and $\alpha\in\mathbb{R}^*$ such that

$$\widetilde{T}_0(E_0)=P^{-1}\begin{pmatrix}1&0\\ \alpha&1\end{pmatrix}P,\quad\text{where}\quad P=\begin{pmatrix}p_{11}&p_{12}\\ p_{21}&p_{22}\end{pmatrix}. \tag{4-35}$$

Thus, by (4-34), $E_0 \in \sigma(H_L)$ if and only if

$$\begin{aligned} 0 &= \left| \begin{pmatrix} a_{k+1}(E_0) & b_{k+1}(E_0) \\ a_k(E_0) & b_k(E_0) \end{pmatrix} (\widetilde{T}_0(E_0))^N \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| \\ &= \left| \begin{pmatrix} a_{k+1}(E_0) & b_{k+1}(E_0) \\ a_k(E_0) & b_k(E_0) \end{pmatrix} P^{-1} \begin{pmatrix} 1 & 0 \\ N\alpha & 1 \end{pmatrix} P \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|; \end{aligned} \tag{4-36}$$

that is,

$$0 = \left| \begin{pmatrix} 1 & 0 \\ N\alpha & 1 \end{pmatrix} P \begin{pmatrix} 1 \\ 0 \end{pmatrix}, P \begin{pmatrix} -b_{k+1}(E_0) \\ a_{k+1}(E_0) \end{pmatrix} \right| = (\det P) a_{k+1}(E_0) - N\alpha p_{11}(-p_{11} b_{k+1}(E_0) + p_{12} a_{k+1}(E_0)).$$

For N large, this expression vanishes if and only if

$$(\det P) a_{k+1}(E_0) = 0 \quad \text{and} \quad \alpha p_{11}(-p_{11} b_{k+1}(E_0) + p_{12} a_{k+1}(E_0)) = 0.$$

Since P is invertible, $|b_{k+1}(E_0)| + |a_{k+1}(E_0)| \neq 0$ and $\alpha \neq 0$, one has $a_{k+1}(E_0) = 0$ and $p_{11} = 0$.

In this case, using $b_{k+1}(E_0) \neq 0$, we can then rewrite the eigenvalue equation (4-36) as

$$0 = \left| (\widetilde{T}_0(E_0))^N \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| = \tilde{t}^{21}_{0,N}(E_0). \tag{4-37}$$

For $E \in \Sigma_{\mathbb{Z}}^{\circ}$ close to E_0, by (4-26) we have

$$t^{21}_{0,N}(E) = \frac{(\rho^N(E) - \rho^{-N}(E)) a^0_{p-1}(E)}{\rho(E) - \rho^{-1}(E)} = \rho^{N-1} \left(\sum_{j=0}^{N-1} \rho^{-2j}(E) \right) a^0_{p-1}(E).$$

As ρ is continuous at E_0 and $\rho^2(E_0) = 1$, taking E to E_0 we get

$$a^0_{p-1}(E_0) = 0.$$

As $\widetilde{T}_0(E_0)$ is not diagonal, this implies $b^0_p(E_0) \neq 0$. This completes the proof of Lemma 4.8. □

Now, pick $E \notin \partial \Sigma_{\mathbb{Z}}$. Then, by Lemma 4.6, the quantization condition (4-34) becomes

$$\begin{vmatrix} \rho^N(E) \dfrac{a^0_p(E) - \rho^{-1}(E)}{\rho(E) - \rho^{-1}(E)} + \rho^{-N}(E) \dfrac{\rho(E) - a^0_p(E)}{\rho(E) - \rho^{-1}(E)} & -b_{k+1}(E) \\ (\rho^N(E) - \rho^{-N}(E)) \dfrac{a^0_{p-1}(E)}{\rho(E) - \rho^{-1}(E)} & a_{k+1}(E) \end{vmatrix} = 0. \tag{4-38}$$

The eigenvalues outside of $\Sigma_{\mathbb{Z}}$. Let us first study the eigenvalues outside $\Sigma_{\mathbb{Z}}$, i.e., in the region $|\Delta(E)| > 2$. If, for $j \in \mathbb{N}$, we define

$$\begin{aligned} \alpha_j(E) &:= a_j(E) \frac{a^0_p(E) - \rho^{-1}(E)}{\rho(E) - \rho^{-1}(E)} + b_j(E) \frac{a^0_{p-1}(E)}{\rho(E) - \rho^{-1}(E)} \\ \text{and} \quad \beta_j(E) &:= a_j(E) \frac{\rho(E) - a^0_p(E)}{\rho(E) - \rho^{-1}(E)} - b_j(E) \frac{a^0_{p-1}(E)}{\rho(E) - \rho^{-1}(E)}, \end{aligned} \tag{4-39}$$

(4-38) can be rewritten as $\beta_{k+1}(E)+\rho^{2N}(E)\alpha_{k+1}(E)=0$; using

$$\alpha_{k+1}(E)+\beta_{k+1}(E)=a_{k+1}(E), \tag{4-40}$$

(4-38) becomes

$$\beta_{k+1}(E)=-\frac{\rho^{2N}(E)}{1-\rho^{2N}(E)}a_{k+1}(E). \tag{4-41}$$

We first show:

Lemma 4.9. *There exists $\eta>0$ such that, for L sufficiently large, $\sigma(H_L)\cap[(\Sigma_{\mathbb{Z}}+[-\eta,\eta])\setminus\Sigma_{\mathbb{Z}}]=\varnothing$.*

Proof. Using (4-39), we rewrite (4-41) as

$$a_{k+1}(E)(\rho(E)-a_p^0(E))-b_{k+1}(E)a_{p-1}^0(E)=\rho^{2N+1}(E)\frac{1-\rho^2(E)}{1-\rho^{2N}(E)}a_{k+1}(E). \tag{4-42}$$

Pick $E_0\in\partial\Sigma_Z$. Then, by our choice for ρ, for $\eta>0$ small we know that, for $E\in[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$, $\rho^2(E)=e^{-c_0\sqrt{|E-E_0|}(1+O(\sqrt{|E-E_0|}))}$. Hence, for $E\in[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$, one has

$$\left|\rho^{2N+1}(E)\frac{1-\rho^2(E)}{1-\rho^{2N}(E)}\right|\lesssim\min\Big(\sqrt{|E-E_0|},\ \frac{1}{N}\Big). \tag{4-43}$$

Thus, if $a_{k+1}(E_0)(\rho(E_0)-a_p^0(E_0))-b_{k+1}(E_0)a_{p-1}^0(E_0)\neq0$, (4-42) has no solution in $[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$ for η small and L sufficiently large.

Let us now assume that $a_{k+1}(E_0)(\rho(E_0)-a_p^0(E_0))-b_{k+1}(E_0)a_{p-1}^0(E_0)=0$.

- If $a_{k+1}(E_0)\neq0$, one computes

$$a_{k+1}(E)(\rho(E)-a_p^0(E))-b_{k+1}(E)a_{p-1}^0(E)=a_{k+1}(E_0)(\rho(E)-\rho(E_0))(1+o(1))$$

and

$$\rho^{2N+1}(E)\frac{1-\rho^2(E)}{1-\rho^{2N}(E)}a_{k+1}(E)=-(\rho(E)-\rho(E_0))a_{k+1}(E_0)\frac{\rho^{2(N+1)}(E)}{1-\rho^{2N}(E)}(1+o(1)).$$

Hence, for $\eta>0$ small and $E\in[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$, the two sides of (4-42) have opposite signs; there is no solution to (4-42) in this interval.

- If $a_{k+1}(E_0)=0$, then $b_{k+1}(E_0)\neq0$, $a_{p-1}^0(E_0)=0$, $\rho(E_0)=a_p^0(E_0)$ and $(a_{p-1}^0)'(E_0)\neq0$; one computes

$$a_{k+1}(E)(\rho(E)-a_p^0(E))-b_{k+1}(E)a_{p-1}^0(E)=-b_{k+1}(E_0)(a_{p-1}^0)'(E_0)(E-E_0)(1+o(1))$$

and, by (4-43), for $\eta>0$ small and $E\in[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$,

$$\left|\rho^{2N+1}(E)\frac{1-\rho^2(E)}{1-\rho^{2N}(E)}a_{k+1}(E)\right|\lesssim|E-E_0|\min\Big(\sqrt{|E-E_0|},\ \frac{1}{N}\Big).$$

Hence, for $\eta>0$ small and $E\in[E_0-\eta,E_0+\eta]\setminus\Sigma_{\mathbb{Z}}$, there is no solution to (4-42) in this interval.

This completes the proof of Lemma 4.9. □

In Lemma 4.8, we saw that, if $E_0\in\partial\Sigma_{\mathbb{Z}}$ satisfies

$$a_{k+1}(E_0)=0\quad\text{and}\quad a_{k+1}(E_0)(\rho(E_0)-a_p^0(E_0))-b_{k+1}(E_0)a_{p-1}^0(E_0)=0,$$

then E_0 is an eigenvalue of H_L for L large.

By Lemma 4.9, it now suffices to consider energies such that $|\underline{\Delta}(E)| > 2+\eta$ for some $\eta > 0$. In this case, we note that the left-hand side in (4-41) is the left-hand side of the first equation in (4-31) (up to the factor $\rho - \rho^{-1}$, which does not vanish outside $\Sigma_{\mathbb{Z}}$). On the other hand, the right-hand side in (4-41) is uniformly exponentially small for large N on $\{E \in \mathbb{R} \mid |\underline{\Delta}(E)| > 2+\eta\}$. Thus, for L large, the solutions to (4-41) are exponentially close to E', which is either an eigenvalue of H_0^+ or one of H_k^-. One distinguishes between the following cases:

(1) If E' is an eigenvalue of H_0^+ but not of H_k^-, then E' is a simple root of the function $E \mapsto \beta_{k+1}(E)$ (see Section 4A2); one has to distinguish two cases depending on whether $a_{k+1}(E')$ vanishes or not. Assume first $a_{k+1}(E') = 0$; then, by (4-28), we know that the eigenvector of H_0^+ actually satisfies the Dirichlet boundary conditions at L; thus, E' is a solution to (4-41), i.e., an eigenvalue of H_L, and (4-28) gives a (nonnormalized) eigenvector.

Assume now that $a_{k+1}(E') \neq 0$; then, by Rouché's theorem, the unique solution to (4-41) close to E' satisfies

$$E - E' = -\frac{\rho^{2N}(E')}{\beta'_{k+1}(E')} a_{k+1}(E')(1 + o(\rho^{2N}(E'))). \tag{4-44}$$

(2) If E' is an eigenvalue of H_k^- but not of H_0^+, mutatis mutandis, the analysis is the same as in point (1).

(3) If E' is an eigenvalue of both H_0^+ and H_k^-, then we are in a resonant tunneling situation. The analysis done in the Appendix shows that, near E', H_L has two eigenvalues, say $E_\pm$, satisfying, for some constant $\alpha > 0$,

$$E_\pm - E' = \pm\alpha\rho^N(E'))\big(1 + O(N\rho(E')^N)\big). \tag{4-45}$$

This completes the proof of the statements of Theorem 4.2 for the eigenvalues outside $\Sigma_{\mathbb{Z}}$.

The eigenvalues inside $\Sigma_{\mathbb{Z}}$. We now study the eigenvalues in the region $\Sigma_{\mathbb{Z}}^\circ$. One can express $\rho(E)$ in terms of the Bloch quasimomentum $\theta_p(E)$ and use $\rho^{-1}(E) = \overline{\rho(E)}$. Notice that, on $\Sigma_{\mathbb{Z}}^\circ$, one has:

- $\operatorname{Im} \rho(E)$ does not vanish.
- The function $E \mapsto \rho(E)$ is real analytic.
- The functions $E \mapsto a_p^0(E)$, $E \mapsto a_{p-1}^0(E)$, $E \mapsto a_{k+1}(E)$ and $E \mapsto b_{k+1}(E)$ are real-valued polynomials.

We prove:

Lemma 4.10. *The function* α_{k+1} *is analytic and does not vanish on* $\Sigma_{\mathbb{Z}}^\circ$.

Proof. Assume that the function α_{k+1} vanishes at a point E_0 in $\Sigma_{\mathbb{Z}}^\circ$.

- If $\rho(E_0) \neq \rho^{-1}(E_0)$, then one has $a_{k+1}(E_0)(a_p^0(E_0) - \rho^{-1}(E_0)) + b_{k+1}(E_0)a_{p-1}^0(E_0) = 0$; as $\rho(E_0) \neq \rho^{-1}(E_0)$ and $E_0 \in \Sigma_{\mathbb{Z}}^\circ$, one has $\rho^{-1}(E_0) = \overline{\rho(E_0)} \notin \mathbb{R}$; thus, for

$$a_{k+1}(E_0)(a_p^0(E_0) - \rho^{-1}(E_0)) - b_{k+1}(E_0)a_{p-1}^0(E_0)$$

to vanish, one needs $a_{k+1}(E_0)=0$ and $a^0_{p-1}(E_0)=0$ (as b_{k+1} and a_{k+1} don't vanish together); this implies that $\rho(E_0)=\pm 1$ and contradicts $\rho(E_0)\neq\rho^{-1}(E_0)$.

• If $\rho(E_0)=\rho^{-1}(E_0)$, such a point E_0 is a simple root of the three functions a^0_{p-1}, $\rho-\rho^{-1}$ and $a^0_p-\rho$ that are analytic near E_0 (see points (1)–(4) in Section 4A1). Moreover, one checks that the derivatives of these functions at that point are respectively real, purely imaginary and neither real nor purely imaginary; for E close to E_0, one has

$$\begin{aligned} a^0_{p-1}(E)&=A(E-E_0)(1+O(E-E_0)),\\ \rho(E)-\rho^{-1}(E)&=2iC(E-E_0)(1+O(E-E_0)),\\ a^0_p(E)-\rho^{-1}(E)&=(B+iC)(E-E_0)(1+O(E-E_0)),\quad \text{where}\quad (A,B,C)\in(\mathbb{R}^*)^3.\end{aligned}\tag{4-46}$$

Now, as a_{k+1} and b_{k+1} are real-valued and can't vanish at the same point, we see that $\alpha_{k+1}(E_0)\neq 0$.

This complete the proof of Lemma 4.10 □

Now, as $L=Np+k$, the characteristic equation (4-38) (valid for $E\in\Sigma^\circ_{\mathbb{Z}}$) becomes

$$\begin{aligned}\rho^{2N}(E)=e^{2iNp\theta_p(E)}&=-\frac{\overline{\alpha_{k+1}(E)}}{\alpha_{k+1}(E)}=-\frac{\beta_{k+1}(E)}{\overline{\beta_{k+1}(E)}}\\ &=\frac{a_{k+1}(E)(\rho(E)-a^0_p(E))-b_{k+1}(E)a^0_{p-1}(E)}{\overline{a_{k+1}(E)(\rho(E)-a^0_p(E))-b_{k+1}(E)a^0_{p-1}(E)}}=:e^{2ih_k(E)}.\end{aligned}\tag{4-47}$$

By Lemma 4.10, the function $E\mapsto h_k(E)$ defined in (4-47) is real analytic on $\Sigma^\circ_{\mathbb{Z}}$. Clearly, as we are inside $\Sigma_{\mathbb{Z}}$, ρ is real only at bands' edges or closed gaps, h_k takes values in $\pi\mathbb{Z}$ only at bands' edges or closed gaps. This implies Theorem 4.2(a). We prove:

Lemma 4.11. *The function h_k can be extended continuously from $\Sigma^\circ_{\mathbb{Z}}$ to $\Sigma_{\mathbb{Z}}$; for $E_0\in\partial\Sigma_{\mathbb{Z}}$, one has*

$$h_k(E_0)\in\begin{cases}\frac{\pi}{2}+\pi\mathbb{Z} & \text{if } a_{k+1}(E_0)\neq 0 \text{ and } a_{k+1}(E_0)(\rho(E_0)-a^0_p(E_0))-b_{k+1}(E_0)a^0_{p-1}(E_0)=0,\\ \pi\mathbb{Z} & \text{if not.}\end{cases}$$

The function $\theta_{p,L}$ is strictly increasing on the bands of $\Sigma_{\mathbb{Z}}$.

Proof. Pick $E_0\in\partial\Sigma_{\mathbb{Z}}$. It suffices to study the behavior of, for $E\in\Sigma_{\mathbb{Z}}$,

$$E\mapsto s(E):=a_{k+1}(E)(\rho(E)-a^0_p(E))-b_{k+1}(E)a^0_{p-1}(E)$$

near E_0 inside $\Sigma_{\mathbb{Z}}$. Write $E=E_0\pm t^2$ for t real and positive; here, the sign $\pm$ depends on whether E_0 is a left or right edge of $\Sigma_{\mathbb{Z}}$ and is chosen so that $E=E_0\pm t^2\in\Sigma^\circ_{\mathbb{Z}}$ for t small.

First, $t\mapsto\rho(E_0\pm t^2)$ is analytic near 0; thus, so is $t\mapsto s(E_0\pm t^2)$. Solving the characteristic equation $\rho^2(E)-\underline{\Delta}(E)\rho(E)+1=0$, one finds

$$\rho(E_0\pm t^2)=\rho(E_0)+iat+bt^2+O(t^3),\quad a\in\mathbb{R}^*,\ b\in\mathbb{R}.$$

Thus,

$$s(E_0\pm t^2)=s(E_0)+ia_{k+1}(E_0)\cdot a\cdot t+c\cdot t^2+O(t^3),$$

where

$$c := a'_{k+1}(E_0)(\rho(E_0) - a^0_p(E_0)) + a_{k+1}(E_0)(b - (a^0_p)'(E_0)) - (b'_{k+1}(E_0)a^0_{p-1}(E_0) + b_{k+1}(E_0)(a^0_{p-1})'(E_0)).$$

Hence:

- If $s(E_0) \neq 0$, then $s(E_0 \pm t^2) = s(E_0) + O(t)$; hence, $h_k(E_0 \pm t^2) = \pi n + O(t)$ for some $n \in \mathbb{Z}$.
- If $s(E_0) = 0$ and $a_{k+1}(E_0) \neq 0$, one has $s(E_0 \pm t^2) = i a_{k+1}(E_0) \cdot a \cdot t + O(t^2)$; thus, $h_k(E_0 \pm t^2) = \frac{\pi}{2} + \pi n + O(t)$ for some $n \in \mathbb{Z}$.
- If $s(E_0) = a_{k+1}(E_0) = 0$, one has $b_{k+1}(E_0) \neq 0$, $a^0_{p-1}(E_0) = 0$, $\rho(E_0) = a^0_p(E_0)$ and $(a^0_{p-1})'(E_0) \neq 0$; thus $s(E_0 \pm t^2) = -b_{k+1}(E_0)(a^0_{p-1})'(E_0)t^2 + 0(t^2)$; hence, $h_k(E_0 \pm t^2) = \pi n + O(t)$ for some $n \in \mathbb{Z}$.

This completes the proof of the statement of Lemma 4.11 on the function h_k.

Let us now control the monotony of $\theta_{p,L}$ (see Theorem 4.2) on the bands of $\Sigma_{\mathbb{Z}}$. It is well known that, keeping the above notations, $\theta_p(E_0 \pm t^2) - \theta_p(E_0) = \pm\alpha t(1 + t g_0(t))$ with $\alpha > 0$. The computations done in the previous paragraph show that $h_k(E_0 \pm t^2) = h_k(E_0) + a t^k(1 + t g_1(t))$, $k \geq 1$. Hence:

- If $k > 1$, we have $\theta_{p,L}(E_0 \pm t^2) - \theta_{p,L}(E_0) = \pm\alpha t(1 + t g_2(t))$.
- If $k = 1$, we have $\theta_{p,L}(E_0 \pm t^2) - \theta_{p,L}(E_0) = (\pm\alpha + a/(L-k))t(1 + t g_2(t))$.

Hence, $\theta_{p,L}$ is strictly increasing inside the band near E_0 for L sufficiently large. Outside a neighborhood of the edges of a band, by analyticity of h_k, as the bands are compact, we have $|\theta'_{p,L} - \theta'_p| \lesssim L^{-1}$. As θ_p is strictly increasing on each band, $\theta_{p,L}$ is also strictly increasing outside a neighborhood of the edges of a band. This completes the proof of Lemma 4.11. □

One proves:

Lemma 4.12. *Let E_0 be a closed gap for $H^{\mathbb{Z}}$* (see Definition 4.5). *Then, for any $L = Np + k$,*

$$E_0 \in \sigma(H_L) \iff h_k(E_0) \in \pi\mathbb{Z} \iff a_{k+1}(E_0) = 0 \iff \alpha_{k+1}(E_0) \in i\mathbb{R}^*. \tag{4-48}$$

Proof. The proof of the first equivalence follows immediately from Definition 4.5 and the quantization condition (4-47); the second follows from (4-39) and the expansions in (4-46); the third follows from Lemma 4.11, (4-39) and (4-47). □

Let us note that, in particular, closed gaps where a_{k+1} vanishes are eigenvalues of H_L for all $L = Np+k$.

Remark 4.13. The characteristic equation (4-47) and the computations done at the end of the proof of Lemma 4.10 show that, for $L = Np + k$ large, an energy E_0 such that $\rho(E_0) = \rho^{-1}(E_0)$ is an eigenvalue of H_L if and only if $a_{k+1}(E_0) = 0$. This is an extension of Lemma 4.8.

In view of the definition and monotony of $\theta_{p,L}$, the quantization condition (4-47) is clearly equivalent to (4-3). This completes the proof Theorem 4.1 on the eigenvalues of H_L. Let us now turn to the computation of the associated eigenfunctions.

4A4. *The Dirichlet eigenfunctions for a truncated periodic potential*: *the proof of Theorem 4.3.* Recall that we assume $L = Np + k$. First, if $(u_l^j)_{l=0}^L$ is an eigenfunction associated to the eigenvalue λ_j, the eigenvalue equation reads

$$\begin{pmatrix} u_{l+1}^j \\ u_l^j \end{pmatrix} = T_l(\lambda_j) \begin{pmatrix} u_l^j \\ u_{l-1}^j \end{pmatrix} \quad \text{for } 0 \le l \le L, \quad \text{where} \quad u_{L+1}^j = u_{-1}^j = 0.$$

To normalize the solution, we assume that $u_0^j = 1$. The coefficients we want to compute are

$$|\varphi_j(L)|^2 = |u_L^j|^2 \left(\sum_{l=0}^L |u_l^j|^2 \right)^{-1} \quad \text{and} \quad |\varphi_j(0)|^2 = \left(\sum_{l=0}^L |u_l^j|^2 \right)^{-1}. \tag{4-49}$$

Fix $l = np + m$. Thus, using the notations of Section 4A3 and the expressions (4-25), (4-26) and (4-23), one computes

$$\begin{pmatrix} u_l^j \\ u_{l-1}^j \end{pmatrix} = T_{m-1,0}(\lambda_j)(\widetilde{T}_0(\lambda_j))^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_m(\lambda_j)\rho^n(\lambda_j) + \beta_m(\lambda_j)\rho^{-n}(\lambda_j) \\ \alpha_{m-1}(\lambda_j)\rho^n(\lambda_j) + \beta_{m-1}(\lambda_j)\rho^{-n}(\lambda_j) \end{pmatrix}, \tag{4-50}$$

where α_m and β_m are as defined in (4-39).

The eigenvectors associated to eigenvalues inside $\Sigma_{\mathbb{Z}}$. As $\rho^{-1}(\lambda_j) = \overline{\rho(\lambda_j)}$, $\beta_m(\lambda_j) = \overline{\alpha_m(\lambda_j)}$ and as the functions $(\alpha_m)_{0 \le m \le p-1}$ do not vanish on $\Sigma_{\mathbb{Z}}^\circ$, we compute

$$|u_{np+m}^j|^2 = 2|\alpha_m(\lambda_j)|^2 \left(1 + \operatorname{Re}\left[\frac{\alpha_m(\lambda_j)}{\overline{\alpha_m(\lambda_j)}} \rho^{2n}(\lambda_j) \right] \right). \tag{4-51}$$

As $L = Np + k$, using the quantization condition (4-47), we obtain that

$$\begin{aligned} &\sum_{l=0}^L |u_l^j|^2 \\ &\quad = 2\sum_{m=0}^k |\alpha_m(\lambda_j)|^2 \left(1 + \operatorname{Re}\left[\frac{\alpha_m(\lambda_j)}{\overline{\alpha_m(\lambda_j)}} \rho^{2N}(\lambda_j) \right] \right) + 2\sum_{m=0}^{p-1} |\alpha_m(\lambda_j)|^2 \sum_{n=0}^{N-1} \left(1 + \operatorname{Re}\left[\frac{\alpha_m(\lambda_j)}{\overline{\alpha_m(\lambda_j)}} \rho^{2n}(\lambda_j) \right] \right) \\ &\quad = Npf(\lambda_j) \left(1 + \frac{1}{Np} \tilde{f}(\lambda_j) \right), \end{aligned} \tag{4-52}$$

where we have defined

$$f(E) := \frac{2}{p} \sum_{m=0}^{p-1} |\alpha_m(E)|^2 \tag{4-53}$$

and, using the quantization condition (4-47), computed

$$\begin{aligned} \tilde{f}(E) := \frac{2}{f(E)} \operatorname{Re}\left[\left(\sum_{m=0}^{p-1} \alpha_m^2(E) \right) \frac{1}{1-\rho^2(E)} \left(1 + \frac{\overline{\alpha_{k+1}(E)}}{\alpha_{k+1}(E)} \right) \right] \\ + \frac{2}{f(E)} \sum_{m=0}^k |\alpha_m(E)|^2 \left(1 - \operatorname{Re}\left[\frac{\alpha_m(E)\overline{\alpha_{k+1}(E)}}{\overline{\alpha_m(E)}\alpha_{k+1}(E)} \right] \right). \end{aligned} \tag{4-54}$$

The function $E \mapsto f(E)$ is real analytic and does not vanish on $\Sigma_{\mathbb{Z}}^{\circ}$.

We prove:

Proposition 4.14. *For E_0, a closed gap, one has $\sum_{m=0}^{p-1} \alpha_m^2(E_0) = 0$.*

Proof. By the definition of (a_j, b_j) — see (4-29) — and that of $\alpha_j(E)$ — see (4-39) — the sequence $(\alpha_j(E))_{j\in\mathbb{Z}}$ satisfies the equation $\alpha_{j+1}+\alpha_{j-1}+(V_j-E)\alpha_j=0$. As $\widetilde{T}_0(E)=T_{p-1}(E)\cdots T_0(E)$, by (4-23), for $j\in\mathbb{Z}$ one has $\alpha_{j+p}(E)=\rho(E)\alpha_j(E)$. Hence, the column vector $A(E)=(\alpha_1(E),\dots,\alpha_p(E))^t$ satisfies

$$(H_\rho - E)A(E) = 0, \quad \text{where} \quad H_\rho = \begin{pmatrix} V_1 & 1 & 0 & \cdots & 0 & \rho(E) \\ 1 & V_2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & V_3 & 1 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & V_{p-1} & 1 \\ \rho^{-1}(E) & 0 & \cdots & 0 & 1 & V_p \end{pmatrix}.$$

Thus, we have

$$\langle (H_\rho - E)A(E), A(E)\rangle_{\mathbb{R}} = 0, \tag{4-55}$$

where $\langle\cdot,\cdot\rangle_{\mathbb{R}}$ denotes the real scalar product over $\mathbb{C}^p$, i.e.,

$$\left\langle \begin{pmatrix} z_1 \\ \vdots \\ z_p \end{pmatrix}, \begin{pmatrix} z_1' \\ \vdots \\ z_p' \end{pmatrix} \right\rangle_{\mathbb{R}} = \sum_{j=1}^{p} z_j z_j'.$$

The functions $E \mapsto A(E)$ and $E \mapsto \rho(E)$ being analytic over $\Sigma_{\mathbb{Z}}^{\circ}$ (see Section 4A1 and Lemma 4.10), one can differentiate (4-55) with respect to E to obtain

$$0 = -\langle A(E), A(E)\rangle_{\mathbb{R}} + (\rho(E)-\rho^{-1}(E))\big(\rho^{-1}(E)\rho'(E)\alpha_1(E)\alpha_p(E) - \alpha_p(E)\alpha_1'(E) + \alpha_1(E)\alpha_p'(E)\big). \tag{4-56}$$

Here we have used the fact that, if H_ρ^t is the transpose of the matrix H_ρ, then

$$H_\rho^t - H_\rho = (\rho(E)-\rho^{-1}(E)) \begin{pmatrix} 0 & \cdots & 0 & -1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

At E_0, a closed gap, one has $\rho(E_0)=\rho^{-1}(E_0)$. Hence, (4-56) implies

$$0 = \langle A(E_0), A(E_0)\rangle_{\mathbb{R}} = \sum_{m=0}^{p-1} \alpha_m^2(E_0).$$

This completes the proof of Proposition 4.14. ☐

In view of (4-54), the function $\tilde{f}$ is real analytic on $\Sigma_{\mathbb{Z}}^{\circ}$; indeed, the only poles of the function $E \mapsto [\rho(E)-\rho^{-1}(E)]^{-1}$ in $\Sigma_{\mathbb{Z}}^{\circ}$ are the closed gaps; they are simple poles of this function and, by Proposition 4.14, the real analytic function $E \mapsto \sum_{m=0}^{p-1} \alpha_m^2(E)$ vanishes at these poles.

Now that we have computed the normalization constant, let us compute the coefficient u_L^j defined in (4-49). As $L = Np+k$, the characteristic equation for λ_j — that is, (4-47) — reads

$$\alpha_{k+1}(\lambda_j)\rho^N(\lambda_j) = -\beta_{k+1}(\lambda_j)\rho^{-N}(\lambda_j) = -\overline{\alpha_{k+1}(\lambda_j)\rho^N(\lambda_j)}. \tag{4-57}$$

Hence, one computes

$$\begin{aligned} u_L^j &= \alpha_k(\lambda_j)\rho^N(\lambda_j) + \overline{\alpha_k(\lambda_j)\rho^N(\lambda_j)} = \rho^N(\lambda_j)\frac{\alpha_k(\lambda_j)\overline{\alpha_{k+1}(\lambda_j)} - \overline{\alpha_k(\lambda_j)}\alpha_{k+1}(\lambda_j)}{\overline{\alpha_{k+1}(\lambda_j)}} \\ &= \frac{-\rho^N(\lambda_j)a_{p-1}^0(\lambda_j)}{(\rho(\lambda_j)-\rho^{-1}(\lambda_j))\overline{\alpha_{k+1}(\lambda_j)}} = \frac{-e^{i[Np\theta_p(\lambda_j)-h_k(\lambda_j)]}a_{p-1}^0(\lambda_j)}{\left|a_{k+1}(\lambda_j)(a_p^0(\lambda_j)-\rho^{-1}(\lambda_j)) + b_{k+1}(\lambda_j)a_{p-1}^0(\lambda_j)\right|} \\ &= \frac{-e^{i\pi j}a_{p-1}^0(\lambda_j)}{\left|a_{k+1}(\lambda_j)(a_p^0(\lambda_j)-\rho^{-1}(\lambda_j)) + b_{k+1}(\lambda_j)a_{p-1}^0(\lambda_j)\right|}, \end{aligned} \tag{4-58}$$

where we have used the quantization condition satisfied by λ_j, the last equality in (4-47), and that

$$\begin{vmatrix} \alpha_{k+1}(\lambda_j) & \alpha_k(\lambda_j) \\ \overline{\alpha_{k+1}(\lambda_j)} & \overline{\alpha_k(\lambda_j)} \end{vmatrix} = \begin{vmatrix} \dfrac{a_{p-1}^0(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} & \dfrac{a_p^0(\lambda_j)-\rho^{-1}(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} \\ -\dfrac{a_{p-1}^0(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} & \dfrac{\rho(\lambda_j)-a_p^0(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} \end{vmatrix} \begin{vmatrix} b_{k+1}(\lambda_j) & b_k(\lambda_j) \\ a_{k+1}(\lambda_j) & a_k(\lambda_j) \end{vmatrix}$$

and

$$\begin{vmatrix} 1 & \dfrac{a_p^0(\lambda_j)-\rho^{-1}(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} \\ -1 & \dfrac{\rho(\lambda_j)-a_p^0(\lambda_j)}{\rho(\lambda_j)-\rho^{-1}(\lambda_j)} \end{vmatrix} = \begin{vmatrix} b_k(\lambda_j) & b_{k+1}(\lambda_j) \\ a_k(\lambda_j) & a_{k+1}(\lambda_j) \end{vmatrix} = 1.$$

Lemma 4.15. *Define the function* $\tilde{f}_k^-(E)$ *by*

$$\tilde{f}_k^-(E) := \frac{|a_{p-1}^0(E)|^2}{|a_{k+1}(E)(a_p^0(E)-\rho^{-1}(E)) + b_{k+1}(E)a_{p-1}^0(E)|^2}.$$

Then the function $\tilde{f}_k^-$ *does not vanish on* $\Sigma_{\mathbb{Z}}^{\circ}$.

Proof. By the definition of α_{k+1}, one has

$$\tilde{f}_k^-(E) = \frac{|a_{p-1}^0(E)|^2}{|\rho(E)-\rho^{-1}(E))|^2|\alpha_{k+1}(E)|^2}.$$

That this expression is well defined and does not vanish on $\Sigma_{\mathbb{Z}}^{\circ}$ follows from Lemma 4.10 and the computations made in the proof thereof. ☐

Plugging (4-58) into this and (4-51) into (4-49), recalling that $u_0^j = 1$, outside the bad closed gaps we obtain (4-4) if

- in addition to (4-53) and (4-54), we set $f_0^+(E) := 1/f(E)$ and $f_k^-(E) = f_0^+(E)\cdot\tilde f_k^-(E)$,
- we remember that the function a_{p-1}^0 only changes sign in the gaps of the spectrum $\Sigma_{\mathbb{Z}}$ (see point (4) in Section 4A1) and set σ_r to be the sign of $-a_{p-1}^0$ on B_r, the r-th band.

By (4-49) and (4-51), we obtain (4-4) using Lemma 4.15. This completes the proof of the statements in Theorem 4.3 on the eigenfunctions of H_L associated to eigenvalues in $\Sigma_{\mathbb{Z}}^\circ$.

Remark 4.16. To complete our study let us also see what happens to the eigenfunctions near the edges of the spectrum. Pick $E_0 \in \partial\Sigma_{\mathbb{Z}}$. One then knows that, for $E \in \Sigma_{\mathbb{Z}}$ with E close to E_0, one has

$$\theta_p(E) - \theta_p(E_0) = a\sqrt{|E-E_0|}(1+o(1)) \tag{4-59}$$

(see the proof of Lemma 4.11).

Let us rewrite $\tilde f$ (see (4-54)) as

$$\tilde f(E) = \frac{2}{f(E)}\left[\sum_{m=0}^{p-1}|\alpha_m(E)|^2\cos\bigl(h_k(E)-2h_{m-1}(E)-p\theta_p(E)\bigr)\right]\frac{\sin(h_k(E))}{\sin(p\theta_p(E))} + \frac{2}{f(E)}\sum_{m=0}^{k}|\alpha_m(E)|^2\bigl(1-\cos\bigl(2(h_k(E)-h_{m-1}(E))\bigr)\bigr). \tag{4-60}$$

Let us first show:

Lemma 4.17. *For any $0 \le m \le p-1$, $E \mapsto 2|\alpha_m(E)|^2/(pf(E))$ can be extended continuously from $\Sigma_{\mathbb{Z}}^\circ$ to $\Sigma_{\mathbb{Z}}$.*

Proof. For $p = 1$ there is nothing to be done as $2|\alpha_m(E)|^2/(pf(E)) \equiv 1$.

For $p \ge 2$, we note that for $0 \le m \le m+1 \le p-1$, as

$$\begin{vmatrix} a_{m+1}(E) & b_{m+1}(E) \\ a_m(E) & b_m(E)\end{vmatrix} = 1$$

by (4-29),

$$\begin{aligned} 0 &= a_{m+1}(E_0)(a_p^0(E_0)-\rho^{-1}(E_0)) + b_{m+1}(E_0)a_{p-1}^0(E_0) \\ &= a_m(E_0)(a_p^0(E_0)-\rho^{-1}(E_0)) + b_m(E_0)a_{p-1}^0(E_0)\end{aligned}$$

if and only if $a_{p-1}^0(E_0) = 0$ (as this implies $a_p^0(E_0) - \rho^{-1}(E_0) = 0$).

Let us assume this is the case. As $p \ge 2$, we know that $\sum_{j=0}^{p-1}|a_j(E_0)|^2 \ne 0$. By (4-46), for at least one $m_0 \in \{0,\dots,p-1\}$ one has $a_{m_0}(E_0) \ne 0$ and $\alpha_{m_0}(E) = bc^{-1}a_{m_0}(E_0) + O(\sqrt{|E-E_0|})$. Hence, $E \mapsto 2|\alpha_m(E)|^2/(pf(E))$ can be continued to E_0, setting

$$\frac{2|\alpha_m(E_0)|^2}{pf(E_0)} = \frac{|a_m(E_0)|^2}{|a_0(E_0)|^2+\cdots+|a_{p-1}(E_0)|^2}.$$

Actually, $f(E)$ can be continued at E_0 by setting

$$f(E_0)=|a_0(E_0)|^2+\cdots+|a_{p-1}(E_0)|^2. \tag{4-61}$$

Let us now assume that $a^0_{p-1}(E_0)\neq 0$. We study the behavior of α_m near E_0. Recall (4-39). Then one has

(1) either $d_m:=a_m(E_0)(a^0_p(E_0)-\rho^{-1}(E_0))+b_m(E_0)a^0_{p-1}(E_0)\neq 0$, in which case, by (4-46), one has $\alpha_m(E)=(d_m c^{-1}/\sqrt{|E-E_0|})(1+o(1))$;

(2) or $d_m=a_m(E_0)(a^0_p(E_0)-\rho^{-1}(E_0))+b_m(E_0)a^0_{p-1}(E_0)=0$, in which case, since for some $A_m\in\mathbb{R}^*$ and $k_m\geq 1$ one has

$$a_m(E)(a^0_p(E)-\rho^{-1}(E_0))+b_m(E)a^0_{p-1}(E)=A_m(E-E_0)^{k_m}(1+o(1)),$$

by (4-46), one can continue α_m to E_0 by setting $\alpha_m(E_0)=\frac{1}{2}a_m(E_0)$.

As $a^0_{p-1}(E_0)\neq 0$, we know that for some $m_0\in\{0,\dots,p-1\}$ we are in case (a). Hence, one has

$$f(E)=\frac{2}{p|E-E_0|}\sum_{m=0}^{p-1}\Bigl|a_m(E_0)(a^0_p(E_0)-\rho^{-1}(E_0))+b_m(E_0)a^0_{p-1}(E_0)\Bigr|^2(1+o(1)) \tag{4-62}$$

and $E\mapsto 2|\alpha_m(E)|^2/(pf(E))$ can be continued to E_0, setting

$$\frac{2|\alpha_m(E_0)|^2}{pf(E_0)}=\frac{|d_m|^2}{|d_0|^2+\cdots+|d_{p-1}|^2}$$

(using the notation introduced in point (a)).

This completes the proof of Lemma 4.17. □

By Lemma 4.11, we know that, for $1\leq k\leq p$ and $E_0\in\partial\Sigma_{\mathbb{Z}}$, one has $2h_k(E_0)\in\pi\mathbb{Z}$. Thus, for $1\leq k\leq p$, $1\leq m\leq p$ and $E_0\in\partial\Sigma_{\mathbb{Z}}$, one has $\cos(h_k(E_0)-2h_{m-1}(E_0)-p\theta_p(E_0))\sin(h_k(E_0))=0$. Using the expansions leading to the proof of Lemma 4.11, one gets

$$\cos(h_k(E)-2h_{m-1}(E)-p\theta_p(E))\sin(h_k(E))=c\sqrt{|E-E_0|}(1+o(1)).$$

Recalling (4-59) and the fact that $p\theta_p(E_0)\in\pi\mathbb{Z}$, Lemma 4.17 implies that $\tilde f$ can be extended continuously up to E_0. Hence, the expansion (4-52) again yields

$$\sum_{l=0}^{L}|u^j_l|^2\asymp Npf(\lambda_j). \tag{4-63}$$

Let us now review the computation (4-58) in this case. We distinguish two cases:

(1) If $a^0_{p-1}(E_0)=0$, then (4-58) and the fact that $a_{k+1}(E_0)\neq 0$ (this case was dealt with in point (1)), yields that, for $|\lambda_j-E_0|$ sufficiently small,

$$|u^j_L|\asymp\sqrt{|\lambda_j-E_0|}.$$

By (4-61) and (4-63), we obtain

$$|\varphi_j(L)|^2 \asymp \frac{|\lambda_j - E_0|}{Np} \quad \text{and} \quad |\varphi_j(0)|^2 \asymp \frac{1}{Np}. \tag{4-64}$$

(2) If $a^0_{p-1}(E_0) \neq 0$, then

(a) if $d_{k+1} \neq 0$ (see case (a) in the proof of Lemma 4.17), by (4-62) and (4-63) one has

$$|\varphi_j(0)|^2 \asymp \frac{|\lambda_j - E_0|}{Np} \quad \text{and} \quad |\varphi_j(L)|^2 \asymp \frac{|\lambda_j - E_0|}{Np}. \tag{4-65}$$

(b) if $d_{k+1} = 0$, by (4-62) and (4-63) one has

$$|\varphi_j(0)|^2 \asymp \frac{|\lambda_j - E_0|}{Np} \quad \text{and} \quad |\varphi_j(L)|^2 \asymp \frac{1}{Np}. \tag{4-66}$$

The eigenvectors associated to eigenvalues outside $\Sigma_{\mathbb{Z}}$. Let us now turn to the eigenfunctions associated to eigenvalues H_L in the gaps of $\Sigma_{\mathbb{Z}}$, i.e., in the region $\{E \mid |\underline{\Delta}(E)| > 2\}$. On $\mathbb{R} \setminus \Sigma_{\mathbb{Z}}$, the eigenvalue $E \mapsto \rho(E)$ is real-valued (recall that we pick it so that $|\rho(E)| < 1$) and so are all the functions $(\alpha_m)_{0 \le m \le p-1}$ and $(\beta_m)_{0 \le m \le p-1}$ (see (4-39)). For $0 \le m \le p-1$, (4-50) yields

$$|u^j_{np+m}|^2 = \alpha^2_m(E)\rho^{2n}(E) + \beta^2_m(E)\rho^{-2n}(E) + 2\alpha_m(E)\beta_m(E). \tag{4-67}$$

As when we studied the eigenvalues of H_L, let us now distinguish the cases when E is close to an eigenvalue of H^+_0 or to an eigenvalue of H^-_k:

(1) Pick E' an eigenvalue of H^+_0 but not an eigenvalue of H^-_k; then recall that $a^0_{p-1}(E') = 0 = a^0_p(E') - \rho(E')$. Thus, for $0 \le m \le p-1$, one has $\beta_m(E') = 0$. Assume that E is close to E'. As E satisfies (4-44), using (4-41), (4-67) becomes

$$|u^j_{np+m}|^2 = \rho^{2n}(E')\left|\alpha_m(E') - \frac{\beta'_m(E')}{\beta'_{k+1}(E')}a_{k+1}(E')[\rho(E') - \rho^{-1}(E')]\rho^{2(N-n)}(E') + O(\rho^{2N}(E))\right|^2$$

for $0 \le m \le p-1$ if $0 \le n \le N-1$ and $0 \le m \le k$ if $n = N$.

Using (4-40), one computes

$$|u^j_{np+m}|^2 = \rho^{2n}(E')\left|a_m(E') - \frac{\beta'_m(E')}{\beta'_{k+1}(E')}a_{k+1}(E')\rho^{2(N-n)}(E') + O(\rho^{2N}(E))\right|^2. \tag{4-68}$$

This yields

$$\sum_{l=0}^{L} |u^j_l|^2 = \sum_{m=0}^{p-1}\sum_{n=0}^{N-1} \rho^{2n}(E')a^2_m(E') + O(N\rho^{2N}(E)) = \frac{1}{1-\rho^2(E')}\sum_{m=0}^{p-1} a^2_m(E') + O(N\rho^{2N}(E)).$$

Moreover, by (4-49), (4-67) and (4-39), as $a^0_{p-1}(E') = 0 = a^0_p(E') - \rho(E')$, we obtain

$$\begin{aligned}|\varphi_j(L)|^2 &= \rho^{2N}(E')\frac{(1-\rho^2(E'))a^2_{k+1}(E')}{[\beta'_{k+1}(E')]^2\sum_{m=0}^{p-1} a^2_m(E')}\begin{vmatrix} \beta'_k(E') & a_k(E') \\ \beta'_{k+1}(E') & a_{k+1}(E')\end{vmatrix}^2 + O(N\rho^{4N}(E)) \\ &= \gamma\rho^{2N}(E') + O(N\rho^{4N}(E)),\end{aligned}$$

where

$$\gamma := \frac{(1-\rho^2(E'))a_{k+1}^2(E')}{[\beta'_{k+1}(E')]^2 \sum_{m=0}^{p-1} a_m^2(E')} \left(\frac{da_{p-1}^0}{dE}(E') \right)^2 > 0.$$

Hence, $|\varphi_j(L)|$ is exponentially small in L (recall $|\rho(E)| < 1$).

(2) If E' is an eigenvalue of H_k^- but not of H_0^+, then, inverting the parts of H_k^- and H_0^+, we see that $|\varphi_j(L)|$ is of order 1. A precise asymptotic can be computed but it won't be needed.

(3) If E' is an eigenvalue of H_0^+ and of H_k^-, the double well analysis done in the Appendix shows that, for normalized eigenvectors, say φ_j, $j = 1, 2$, associated to the two eigenvalues of H_L close to E', the four coefficients $|\varphi_j(0)|$ and $|\varphi_j(L)|$, $j = 1, 2$, are of order 1. Again, precise asymptotics can be computed but won't be needed.

This completes the description of the eigenfunctions given by Theorem 4.3 and completes the proof of this result.

5. Resonances in the periodic case

We are now in the state to prove the results stated in Section 1B. We first study the function $E \mapsto S_L(E)$ and $E \mapsto \Gamma_L(E)$ in the complex strip $I + i(-\infty, 0)$ for $I \subset \Sigma_{\mathbb{Z}}^{\circ}$.

5A. *The matrix Γ_L in the periodic case.* Using Theorem 4.2, we first prove:

Theorem 5.1. *Fix $I \subset \Sigma_{\mathbb{Z}}^{\circ}$ a compact interval. There exists $\varepsilon_I > 0$ and $\sigma_I \in \{+1, -1\}$ such that, for any $N \geq 0$, there exists $C_N > 0$ such that, for L sufficiently large with $L \equiv k$* mod p, *one has*

$$\sup_{\substack{\operatorname{Re} E \in I \\ -\varepsilon_I < \operatorname{Im} E < 0}} |\Gamma_L(E) - \Gamma_L^{\mathrm{eff}}(E)| \leq C_N L^{-N}, \tag{5-1}$$

where

$$\Gamma_L^{\mathrm{eff}}(E) = -\frac{\theta'_p(E)}{\sin u_L(E)} \begin{pmatrix} e^{-iu_L(E)} f_k^-(E) & \sigma_I \sqrt{f_k^-(E) f_0^+(E)} \\ \sigma_I \sqrt{f_k^-(E) f_0^+(E)} & e^{-iu_L(E)} f_0^+(E) \end{pmatrix} + \begin{pmatrix} \int_{\mathbb{R}} 1/(\lambda - E)\, dN_k^-(\lambda) & 0 \\ 0 & \int_{\mathbb{R}} 1/(\lambda - E)\, dN_0^+(\lambda) \end{pmatrix} \tag{5-2}$$

and $u_L(E) := (L-k)\theta_{p,L}(E)$ (see (4-2)*),*

The sign σ_I only depends on the spectral band containing I.

Deeper in the lower half-plane, we obtain the following simpler estimate:

Theorem 5.2. *There exists $C > 0$ such that, for any $\varepsilon > 0$ and $L \geq 1$ sufficiently large with $L = Np + k$, one has*

$$\sup_{\substack{\operatorname{Re} E \in I \\ \operatorname{Im} E < -\varepsilon}} \left| \Gamma_L(E) - \begin{pmatrix} \int_{\mathbb{R}} 1/(\lambda - E)\, dN_k^-(\lambda) & 0 \\ 0 & \int_{\mathbb{R}} 1/(\lambda - E)\, dN_0^+(\lambda) \end{pmatrix} \right| \leq C\varepsilon^{-2} e^{-\varepsilon L/C}. \tag{5-3}$$

In Section 5B, the approximations (5-1) and (5-3) will be used to prove Theorems 1.5, 1.7 and 1.10.

Let us note that, as $\cot z = i + O(e^{-2i \operatorname{Im} z})$, for $\varepsilon \in (0, \varepsilon_I)$ the asymptotics given by Theorems 5.1 and 5.2 coincide in the region $\{\operatorname{Re} E \in I,\ \operatorname{Im} E \in (-\varepsilon_I, -\varepsilon)\}$; indeed, one has

$$\sup_{\substack{\operatorname{Re} E \in I \\ -\varepsilon_I < \operatorname{Im} E < -\varepsilon}} \left\| \frac{\theta_p'(E)}{\sin u_L(E)} \begin{pmatrix} e^{-iu_L(E)} f_k^-(E) & \sigma_I \sqrt{f_k^-(E) f_0^+(E)} \\ \sigma_I \sqrt{f_k^-(E) f_0^+(E)} & e^{-iu_L(E)} f_0^+(E) \end{pmatrix} \right\| \le e^{-\varepsilon L/C}.$$

Let us now turn to the proofs of Theorems 5.1 and 5.2.

5A1. *The proof of Theorem 5.1.* To prove Theorem 5.1, we split the sum $S_L(E)$ into two parts, one containing the Dirichlet eigenvalues “close” to $\operatorname{Re} E$, the other containing those “far” from $\operatorname{Re} E$. By “far”, we mean that the distance to $\operatorname{Re} E$ is bounded from below by a small constant independent of L. The “close” eigenvalues are then described by Theorem 4.2. For the “far” eigenvalues, the strong resolvent convergence of H_L to H_0^+, that of $\widetilde{H}_L$ to H_k^- (see Remark 1.6), and Combes–Thomas estimates enable us to compute the limit and to show that the prelimit and the limit are $O(L^{-\infty})$ close to each other. For the “close” eigenvalues, the sum occurring in (2-9), the definition of Γ_L, is a Riemann sum. We use the Poisson summation formula to obtain a precise approximation.

As I is a compact interval in $\Sigma_{\mathbb{Z}}^\circ$, we pick $\varepsilon > 0$ such that, for $E \in I$, one has $[E - 6\varepsilon, E + 6\varepsilon] \subset \Sigma_{\mathbb{Z}}^\circ$. Let $\chi \in \mathscr{C}_0^\infty(\mathbb{R})$ be a nonnegative cut-off function such that $\chi \equiv 1$ on $[-4\varepsilon, 4\varepsilon]$ and $\chi \equiv 0$ outside $[-5\varepsilon, 5\varepsilon]$. For $E \in I$, define $\chi_E(\cdot) = \chi(\cdot - E)$.

We first give the asymptotic for the sum over the Dirichlet eigenvalues far from $\operatorname{Re} E$. We prove:

Lemma 5.3. *For any $N > 1$, there exists $C_N > 0$ such that, for L sufficiently large with $L \equiv k \bmod p$, one has*

$$\sup_{E \in \mathbb{C}} \left| \sum_{j=1}^{L} \frac{1 - \chi_{\operatorname{Re} E}(\lambda_j)}{\lambda_j - E} \begin{pmatrix} |\varphi_j(L)|^2 & \overline{\varphi_j(0)}\varphi_j(L) \\ \varphi_j(0)\overline{\varphi_j(L)} & |\varphi_j(0)|^2 \end{pmatrix} - \widetilde{M}(E) \right| \le C_N L^{-N}, \tag{5-4}$$

where

$$\widetilde{M}(E) := \begin{pmatrix} \int_{\mathbb{R}} (1 - \chi_{\operatorname{Re} E})(\lambda)/(\lambda - E)\, dN_k^-(\lambda) & 0 \\ 0 & \int_{\mathbb{R}} (1 - \chi_{\operatorname{Re} E})(\lambda)/(\lambda - E)\, dN_0^+(\lambda) \end{pmatrix}. \tag{5-5}$$

Proof of Lemma 5.3. Recall (see Theorem 2.2) that H_L is the operator H_0^+ restricted to $[\![0, L]\!]$ with Dirichlet boundary condition at L; as $L \equiv k \bmod p$, it is unitarily equivalent to the operator H_k^- restricted to $[\![-L, 0]\!]$ with Dirichlet boundary condition at $-L$ (see Remark 1.6).

Pick $\tilde{\chi} \in \mathscr{C}_0^\infty$ such that $\tilde{\chi} \equiv 1$ on $\sigma(H_0^+) \cup \sigma(H_k^-)$. First, we compute

$$\begin{aligned}\sum_{j=0}^{L} (1 - \chi_{\operatorname{Re} E})(\lambda_j) \frac{|\varphi_j(0)|^2}{\lambda_j - E} - \int_{\mathbb{R}} (1 - \chi_{\operatorname{Re} E})(\lambda)\, \frac{dN_0^+(\lambda)}{\lambda - E} \\ = \langle \delta_0, [\tilde{\chi}(1 - \chi_{\operatorname{Re} E})](H_L)(H_L - E)^{-1}\delta_0 \rangle - \langle \delta_0, [\tilde{\chi}(1 - \chi_{\operatorname{Re} E})](H_0^+)(H_0^+ - E)^{-1}\delta_0 \rangle,\end{aligned}$$

$$\sum_{j=0}^{L}(1-\chi_{\operatorname{Re}E})(\lambda_j)\frac{|\varphi_j(L)|^2}{\lambda_j-E}-\int_{\mathbb{R}}(1-\chi_{\operatorname{Re}E})(\lambda)\,\frac{dN_k^-(\lambda)}{\lambda-E}$$
$$=\big\langle\delta_L,[\tilde\chi(1-\chi_{\operatorname{Re}E})](H_L)(H_L-E)^{-1}\delta_L\big\rangle-\big\langle\delta_L,[\tilde\chi(1-\chi_{\operatorname{Re}E})](H_k^-)(H_k^--E)^{-1}\delta_L\big\rangle,$$

and

$$\sum_{j=0}^{L}(1-\chi_{\operatorname{Re}E})(\lambda_j)\frac{\varphi_j(L)\overline{\varphi_j(0)}}{\lambda_j-E}=\big\langle\delta_L,[\tilde\chi(1-\chi_{\operatorname{Re}E})](H_L)(H_L-E)^{-1}\delta_0\big\rangle.$$

By the definition of $\chi_{\operatorname{Re}E}$, the function $\lambda\mapsto(\lambda-E)^{-1}\tilde\chi(\lambda)(1-\chi_{\operatorname{Re}E})(\lambda)$ is $\mathscr{C}_0^\infty$ on $\mathbb{R}$; moreover, its seminorms (see (4-14)) are bounded uniformly in $E\in\mathbb{C}$. Thus there exists an almost analytic extension of $[\tilde\chi(1-\chi_{\operatorname{Re}E})](\cdot)(\cdot-E)^{-1}$ such that, uniformly in E, one has (4-14).

In the same way as we obtained (4-16), we obtain

$$\big|\big\langle\delta_L,[(\widetilde{H}_L-z)^{-1}-(H_k^--z)^{-1}]\delta_L\big\rangle\big|+\big|\big\langle\delta_0,[(H_L-z)^{-1}-(H_0^+-z)^{-1}]\delta_0\big\rangle\big|+|\langle\delta_0,(H_L-z)^{-1}\delta_L\rangle|$$
$$\le\frac{C}{|\operatorname{Im}z|^2}e^{-L|\operatorname{Im}z|/C}.\tag{5-6}$$

Plugging (5-6) into (4-15) and using (4-14) for $[\tilde\chi(1-\chi_{\operatorname{Re}E})](\cdot)(\cdot-E)^{-1}$, we get

$$\sup_{\substack{L\ge1\\ L\equiv k\bmod p}}L^K\left|\sum_{j=0}^{L}(1-\chi_{\operatorname{Re}E})(\lambda_j)\frac{|\varphi_j(0)|^2}{\lambda_j-E}-\int_{\mathbb{R}}(1-\chi_{\operatorname{Re}E})(\lambda)\,\frac{dN_0^+(\lambda)}{\lambda-E}\right|<+\infty\quad\text{for all }K\in\mathbb{N}.$$

This entails (5-4) and completes the proof of Lemma 5.3. □

Let us now estimate the part of $\Gamma_L(E)$ associated to the Dirichlet eigenvalues close to Re E. Define

$$\Gamma_L^\chi(E)=\sum_{j=1}^{L}\frac{\chi_{\operatorname{Re}E}(\lambda_j)}{\lambda_j-E}\begin{pmatrix}|\varphi_j(L)|^2&\overline{\varphi_j(0)}\varphi_j(L)\\ \varphi_j(0)\overline{\varphi_j(L)}&|\varphi_j(0)|^2\end{pmatrix}.\tag{5-7}$$

We prove:

Lemma 5.4. *There exists $\varepsilon>0$ such that, for $N\ge1$, there exists C_N such that, for L sufficiently large with $L\equiv k$ mod p, one has*

$$\sup_{\substack{\operatorname{Re}E\in I\\ -\varepsilon<\operatorname{Im}E<0}}|\Gamma_L^\chi(E)-\Gamma_L^{\mathrm{eff}}(E)+\widetilde{M}(E)|\le C_NL^{-N},$$

where $\widetilde{M}$ is as defined in (5-5).

Clearly Lemmas 5.3 and 5.4 immediately yield Theorem 5.1.

Proof of Lemma 5.4. Recall that the quasimomentum θ_p defines a real analytic one-to-one monotonic map from the interior of each band of spectrum onto the set $(0,\pi)$, $(-\pi,0)$ or $(-\pi,\pi)$ (depending on the spectral band containing $I+[-4\varepsilon,4\varepsilon]$, where $\varepsilon>0$ has been fixed above) (see, e.g., [Teschl 2000]). Moreover, the derivative θ_p' is positive in the interior of a spectral band. Thus, for L sufficiently large, the real part of the derivative $\theta'_{p,L}$ (see (4-2)) is positive $I+[-3\varepsilon, 3\varepsilon]$ and $\theta_{p,L}$ is real analytic one-to-one on a complex neighborhood of $(I+[-3\varepsilon, 3\varepsilon])+i[-3\varepsilon, 3\varepsilon]$ (possibly at the expense of reducing ε somewhat).

By (2-9), (4-8) and Theorem 4.2, one may write

$$\Gamma_L^\chi(E) = \frac{1}{L-k}\sum_{j\in\mathbb{Z}} \frac{\chi_{\operatorname{Re} E}\big(\theta_{p,L}^{-1}(\pi j/(L-k))\big)}{\theta_{p,L}^{-1}(\pi j/(L-k))-E} M\left(\theta_{p,L}^{-1}\left(\frac{\pi j}{L-k}\right)\right), \tag{5-8}$$

where

$$M(\lambda) := \begin{pmatrix} f_{k,L}(\lambda) & \sigma_I e^{i(L-k)\theta_{p,L}(\lambda)}\sqrt{f_{k,L}(\lambda) f_{0,L}(\lambda)} \\ \sigma_I e^{i(L-k)\theta_{p,L}(\lambda)}\sqrt{f_{k,L}(\lambda) f_{0,L}(\lambda)} & f_{0,L}(\lambda) \end{pmatrix} \tag{5-9}$$

and the matrix M is analytic in the rectangle $(I+[-3\varepsilon, 3\varepsilon])+i[-3\varepsilon, 3\varepsilon]$. Thus, the Poisson formula tells us that

$$\begin{aligned} \Gamma_L^\chi(E) &= \frac{1}{L-k}\sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} e^{-2i\pi jx} \frac{\chi_{\operatorname{Re} E}\big(\theta_{p,L}^{-1}(\pi x/(L-k))\big)}{\theta_{p,L}^{-1}(\pi x/(L-k))-E} M\left(\theta_{p,L}^{-1}\left(\frac{\pi x}{L-k}\right)\right) dx \\ &= \sum_{j\in\mathbb{Z}} \frac{1}{\pi}\int_{\mathbb{R}} e^{-2ij(L-k)\theta_{p,L}(\lambda)} \frac{\chi_{\operatorname{Re} E}(\lambda)}{\lambda - E}\theta'_{p,L}(\lambda) M(\lambda)\, d\lambda \\ &= \sum_{j\in\mathbb{Z}} \frac{1}{\pi}\int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\lambda)\, d\lambda \end{aligned} \tag{5-10}$$

by the definition of $\chi_{\operatorname{Re} E}$; here, we have set

$$M_{j,\chi}(E,\lambda,\beta) := e^{-2ij(L-k)\theta_{p,L}(\beta+\operatorname{Re} E)} \frac{\chi(\lambda)}{\beta - i\operatorname{Im} E}\theta'_{p,L}(\beta+\operatorname{Re} E) M(\beta+\operatorname{Re} E).$$

Let us now study the individual terms in the last sum in (5-10). Recall that, on $[-4\varepsilon, 4\varepsilon]$, χ is identically 1 and that $\lambda \mapsto \theta_{p,L}(\lambda+\operatorname{Re} E)$ and $\lambda \mapsto M(\lambda)$ are analytic in $(I+[-3\varepsilon, 3\varepsilon])+i[-3\varepsilon, 3\varepsilon]$; moreover, by (4-3), for some $\delta > 0$ one has

$$\liminf_{L\to+\infty} \inf_{\lambda\in[-4\varepsilon,4\varepsilon]} \theta'_{p,L}(\lambda+\operatorname{Re} E) \ge \liminf_{L\to+\infty} \inf_{E\in I} \theta'_{p,L}(E) \ge \delta. \tag{5-11}$$

Recall also that $\operatorname{Im} E < 0$. Consider $\tilde\chi : \mathbb{R} \to [0,1]$ smooth such that $\tilde\chi = 1$ on $[-2\varepsilon, 2\varepsilon]$ and $\tilde\chi = 0$ outside $[-3\varepsilon, 3\varepsilon]$.

In the complex plane, consider the paths $\gamma_\pm : R \to \mathbb{C}$ defined by

$$\gamma_\pm(\lambda) = \lambda \pm 2i\varepsilon\tilde\chi(\lambda).$$

As $-\varepsilon \le \operatorname{Im} E < 0$, by contour deformation we have

$$\begin{aligned} \int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\lambda)\, d\lambda &= \int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\gamma_+(\lambda))\, d\lambda \\ &= -2i\pi e^{-2ij(L-k)\theta_{p,L}(E)}\theta'_{p,L}(E)M(E) + \int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\gamma_-(\lambda))\, d\lambda. \end{aligned}$$

We then estimate:

• For $j<0$, using a nonstationary phase argument since the integrand is the product of a smooth function with an rapidly oscillating function (using $|j|(L-k)$ as the large parameter), one then estimates

$$\int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\gamma_+(\lambda))\,d\lambda = O\big((|j|L)^{-\infty}\big).$$

The phase function is complex but its real part is nonpositive as $\operatorname{Im}\theta_{p,L}(\gamma_+(\cdot)+\operatorname{Re}E)\geq 0$ on the support of χ (by (5-11)). Note that the off-diagonal terms of $M(\lambda)$ also carry a rapidly oscillating exponential (see (5-9)) but it clearly does not suffice to counter the main one.

• In the same way, for $j>0$, one has

$$\int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\gamma_-(\lambda))\,d\lambda = O((|j|L)^{-\infty}).$$

Thus, we compute

$$\int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\lambda)\,d\lambda = O\big((|j|L)^{-\infty}\big) \quad \text{for } j<0, \tag{5-12}$$

$$\int_{\mathbb{R}} M_{j,\chi}(E,\lambda,\lambda)\,d\lambda = -2i\pi e^{-2ij(L-k)\theta_{p,L}(E)}\theta'_{p,L}(E)M(E) + O((|j|L)^{-\infty}) \quad \text{for } j>0. \tag{5-13}$$

Finally, for $j=0$, the contour deformation along γ_+ yields

$$\begin{aligned}\int_{\mathbb{R}} \frac{\chi(\lambda)}{\lambda-i\operatorname{Im}E}M(\lambda+\operatorname{Re}E)\,d\lambda &= \int_{\mathbb{R}} \frac{\chi_{\operatorname{Re}E}(\lambda)}{\lambda-E}\theta'_{p,L}(\lambda)\begin{pmatrix} f_{k,L}(\lambda) & 0\\ 0 & f_{0,L}(\lambda)\end{pmatrix}d\lambda + O(L^{-\infty})\\ &= \int_{\mathbb{R}} \frac{\chi_{\operatorname{Re}E}(\lambda)}{\lambda-E}\begin{pmatrix} dN_k^-(\lambda) & 0\\ 0 & dN_0^+(\lambda)\end{pmatrix} + O(L^{-\infty})\end{aligned}$$

by Corollary 4.4.

Plugging this, (5-12) and (5-13) into (5-10) and computing the geometric sum immediately yields the asymptotic expansion (where the remainder term is uniform on the rectangle $I+i[-\varepsilon,0)$)

$$\begin{aligned}\Gamma_L^{\chi}(E) &= -2i\sum_{j>0} e^{-2ij(L-k)\theta_{p,L}(E)}\theta'_{p,L}(E)M(E) + \int_{\mathbb{R}} \frac{\chi_{\operatorname{Re}E}(\lambda)}{\lambda-E}\begin{pmatrix} dN_k^-(\lambda) & 0\\ 0 & dN_0^+(\lambda)\end{pmatrix} + O(L^{-\infty})\\ &= \frac{-e^{-i(L-k)\theta_{p,L}(E)}}{\sin((L-k)\theta_{p,L}(E))}\theta'_{p,L}(E)M(E) + \int_{\mathbb{R}} \frac{\chi_{\operatorname{Re}E}(\lambda)}{\lambda-E}\begin{pmatrix} dN_k^-(\lambda) & 0\\ 0 & dN_0^+(\lambda)\end{pmatrix} + O(L^{-\infty}).\end{aligned} \tag{5-14}$$

This completes the proof of Lemma 5.4. □

5A2. *The proof of Theorem 5.2.* To prove (5-1), for $\operatorname{Im}E<-\varepsilon$ it suffices to write

$$\begin{aligned}\sum_{j=0}^{L}\frac{|\varphi_j(0)|^2}{\lambda_j-E} - \int_{\mathbb{R}}\frac{dN_0^+(\lambda)}{\lambda-E} &= \langle\delta_0,(H_L-E)^{-1}\delta_0\rangle - \langle\delta_0,(H_0^+-E)^{-1}\delta_0\rangle\\ &= \langle\delta_0,(H_L-E)^{-1}\delta_L\rangle\langle\delta_{L+1},(H_0^+-E)^{-1}\delta_0\rangle\end{aligned}$$

and

$$\sum_{j=0}^{L}\frac{|\varphi_j(L)|^2}{\lambda_j-E}-\int_{\mathbb{R}}\frac{dN_k^-(\lambda)}{\lambda-E}=\langle\delta_0,(H_L-E)^{-1}\delta_L\rangle\langle\delta_{L+1},(H_k^--E)^{-1}\delta_0\rangle,$$

$$\sum_{j=0}^{L}\frac{\varphi_j(L)\overline{\varphi_j(0)}}{\lambda_j-E}=\langle\delta_L,(H_L-E)^{-1}\delta_0\rangle,$$

and to use the Combes–Thomas estimate (5-6). This completes the proof of Theorem 5.2.

5B. ***The proofs of Theorems 1.5, 1.7 and 1.10.*** We will now use Theorems 5.1 and 5.2 to prove Theorems 1.5, 1.7 and 1.10.

5B1. *The proof of Theorem 1.5.* The first statement of Theorem 1.5 is an immediate consequence of the characteristic equations for the resonances (2-4) and (2-8) and the description of the eigenvalues of H_L given in Theorem 4.2.

When $\bullet=\mathbb{N}$, i.e., for the operator on the half-line, if $I\subset(-2,2)$ does not meet $\Sigma_{\mathbb{N}}$, there exists $C>0$ such that, for L sufficiently large, $\mathrm{dist}(I,\sigma(H_L))>1/C$. Thus, on the set $I-i[0,+\infty)$, one has $\operatorname{Im}S_L(E)\le\operatorname{Im}E/C$. Since on I one has $\operatorname{Im}\theta_p(E)>1/C$ (see Section 2), the characteristic equation (2-4) admits a solution E such that $\operatorname{Re}E\in I$ only if $\operatorname{Im}E<1/C^2$. This completes the proof of Theorem 1.5(1) for $\bullet=\mathbb{N}$.

For $\bullet=\mathbb{Z}$, i.e., to study (2-8), one reasons in the same way except that one replaces the study of $S_L(E)$ by that of $\langle\Gamma_L(E)u,u\rangle$ for u an arbitrary vector in $\mathbb{C}^2$ of unit length. This completes the proof of Theorem 1.5(1).

Point (3a) is an immediate consequence of Theorems 3.3 and 3.2 and the description of the eigenvalues of H_L outside $\Sigma_{\mathbb{Z}}$. Notice that, in the present case, d_j in Theorems 3.3 and 3.2 is bounded from below by a constant independent of L, and $a_j^\bullet$ is exponentially small and described by Theorem 4.2.

Point (3b) is an immediate consequence of the description of the eigenvalues of H_L outside $\Sigma_{\mathbb{Z}}$ in Theorems 5.2(2) and 3.1. Indeed, in the present case, d_j and $a_j^\bullet$ are both of order 1; thus, Theorem 3.1 guarantees, around the common eigenvalue for H_k^- and H_0^+, a rectangle of width of order 1 free of resonances.

Let us now turn to the proof of point (2). We first prove the following corollary of Theorem 5.1:

Corollary 5.5. *Fix $I\subset\Sigma_{\mathbb{Z}}^\circ$ compact. There exists $\eta_0>0$ such that, for L sufficiently large, one has*

$$\min_{\substack{\operatorname{Re}E\in I\\ \operatorname{Im}E\in[-\eta_0/L,0)}}|S_L(E)+e^{-i\theta(E)}|\ge\eta_0\quad\text{and}\quad\min_{\substack{\operatorname{Re}E\in I\\ \operatorname{Im}E\in[-\eta_0/L,0)}}|\det(\Gamma_L(E)+e^{-i\theta(E)})|\ge\eta_0. \tag{5-15}$$

Clearly, Corollary 5.5 implies that neither (2-4) nor (2-8) can have a solution in $I+i]-\eta_0/L,0]$. This proves Theorem 1.5(2).

Before proving Corollary 5.5, we first prove Propositions 5.7 and 5.8, as these will be used in the proof of Corollary 5.5.

5B2. *Results on the auxiliary functions defined in Section 1B2.* Recall that N_k^- is defined in Section 1B2. We prove:

Proposition 5.6. *For $k \in \{0, \dots, p-1\}$, dN_k^- is a positive measure that is absolutely continuous on $\Sigma_{\mathbb{Z}}$. Moreover, its density, say $E \mapsto n_k^-(E)$, is real analytic on $\Sigma_{\mathbb{Z}}^\circ$ and there exists $f_k^- : \Sigma_{\mathbb{Z}}^\circ \to \mathbb{R}$ a positive real analytic function such that, on $\Sigma_{\mathbb{Z}}^\circ$, one has $n_k^-(E) = f_k^-(E) n(E)$.*

Proof. Proposition 5.6 is an immediate consequence of Theorems 5.1 and 5.2 and Corollary 4.4. □

For Ξ_k^- defined in (1-5), we prove:

Proposition 5.7. *Ξ_k^- vanishes identically if and only if $V \equiv 0$, i.e., V vanishes identically. Moreover, if $V \not\equiv 0$ then there exists $\xi_k^- \neq 0$ and $\alpha_k^- \in \{2, 3, \dots\}$ such that $\Xi_k^-(E) \sim \xi_k^- E^{-\alpha_k^-}$ as $|E| \to \infty$, $\operatorname{Im} E < 0$.*

Proof. We will do the proofs for the function Ξ_k^-. Proposition 5.7 is an immediate consequence of the fact that, in the lower half-plane, the function $E \mapsto -e^{-i \arccos(E/2)} = -\frac{1}{2}E - \sqrt{\frac{1}{4}E^2 - 1}$ (i.e., the choice of it defined above) is equal to the Stieltjes (or Borel) transform of the spectral measure associated to the Dirichlet Laplacian on $\mathbb{N}$ and the vector δ_0; this follows from a direct computation (see Remark 2.1 and (2-2) for $n = 0$). Now, if one lets W be the symmetric of $\tau_k V$ with respect to 0, the spectral measure dN_k^- is also the spectral measure of the Schrödinger operator $H_k = -\Delta + W$ on $\mathbb{N}$ associated to δ_0. The equality of the Borel transforms implies the equality of the measures but δ_0 is cyclic for both operators, so the operators have equal spectral measures. This implies that the two operators are equal and, thus, the symmetric of $\tau_k V$ has to vanish identically on $\mathbb{N}$. As V is periodic, V must vanish identically.

As for the second point, if the function Ξ_k^- were to vanish to infinite order at $E = -i\infty$, as each of the terms $\int_{\mathbb{R}} 1/(\lambda - E)\, dN_k^-(\lambda)$ and $-\frac{1}{2}E - \sqrt{\frac{1}{4}E^2 - 1}$ admits an infinite asymptotic expansion in powers of E^{-1}, these two expansions would be equal. The n-th coefficient of these expansions are the n-th moments of the spectral measures of H_k and $-\Delta_0^+$, respectively (associated to the cyclic vector δ_0). So these moments would coincide and, thus, the spectral measures would coincide. One concludes as above. □

For $c^\bullet$ defined in (1-6) and (1-7), we prove:

Proposition 5.8. *Pick $\bullet \in \{\mathbb{N}, \mathbb{Z}\}$. Let $I \subset (-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$ be a compact interval.*

There exists a neighborhood of I such that, in this neighborhood, the function $E \mapsto c^\bullet(E)$ is analytic and has a positive imaginary part.

The function $c^{\mathbb{N}}$ (resp. $c^{\mathbb{Z}}$) takes the value i only at the zeros of Ξ_k^- (resp. $\Xi_k^- \Xi_0^+$).

Proof. On $\{\operatorname{Im} E < 0\}$, define the functions

$$g_k^-(E) := i + \frac{\Xi_k^-(E)}{\pi n_k^-(E)} = \frac{1}{\pi n_k^-(E)}(S_k^-(E) + e^{-i \arccos(E/2)}), \tag{5-16}$$

$$g_0^+(E) := i + \frac{\Xi_0^+(E)}{\pi n_0^+(E)} = \frac{1}{\pi n_0^+(E)}(S_0^+(E) + e^{-i \arccos(E/2)}). \tag{5-17}$$

First, the analyticity of g_k^- and g_0^+ is clear; indeed, all the functions involved are analytic and the functions n_0^+ and n_k^- stay positive on $\Sigma_{\mathbb{Z}}^\circ$. Moreover, these functions can be analytically continued through

$(-2,2)\cap\Sigma_Z^\circ$. By (1-4), for E real one has $\operatorname{Im} g_k^-(E)=\operatorname{Im} g_0^+(E)=\operatorname{Im} e^{-i\theta(E)}$, which is positive (see Section 2). Thus the functions $E\mapsto g_k^-(E)$ and $E\mapsto g_0^+(E)$ do not vanish on I. Moreover, as

$$\frac{g_0^+(E)g_k^-(E)-1}{g_0^+(E)+g_k^-(E)}=-\frac{1}{g_0^+(E)+g_k^-(E)}+\frac{1}{1/g_0^+(E)+1/g_k^-(E)},\tag{5-18}$$

this function has a positive imaginary part on I.

This proves the first two properties of $c^\bullet$ stated in Proposition 5.8. By the very definition of $c^\bullet$ and g_k^-, the last property stated in Proposition 5.8 is obviously satisfied in the case of the half-line; for the full line, i.e., if $\bullet=\mathbb{Z}$, the last property is a consequence of the computation

$$\begin{aligned}c^{\mathbb{Z}}(E)-i&=\frac{g_0^+(E)g_k^-(E)-1}{g_0^+(E)+g_k^-(E)}-i=\frac{(g_0^+(E)-i)(g_k^-(E)-i)}{g_0^+(E)+g_k^-(E)}\\&=\frac{\Xi_0^+(E)\Xi_k^-(E)}{2i\pi^2n_0^+(E)n_k^-(E)+\pi n_k^-(E)\Xi_0^+(E)+\pi n_0^+(E)\Xi_k^-(E)}.\end{aligned}\tag{5-19}$$

This completes the proof of Proposition 5.8. □

5B3. *The proof of Corollary 5.5.* In view of Theorem 5.1, to obtain (5-15) it suffices to prove that there exists $\eta_0>0$ such that, for L sufficiently large, one has

$$\min_{\substack{\operatorname{Re}E\in I\\ \operatorname{Im}E\in[-\eta_0/L,0)}}\left|\frac{\theta'_{p,L}(E)f_k^-(E)e^{-iu_L(E)}}{\sin u_L(E)}-\int_{\mathbb{R}}\frac{dN_k^-(\lambda)}{\lambda-E}-e^{-i\theta(E)}\right|\ge\eta_0,$$

where $u_L(E):=(L-k)\theta_{p,L}(E)$.

We compute

$$\frac{\theta'_{p,L}(E)f_k^-(E)e^{-iu_L(E)}}{\sin u_L(E)}-\int_{\mathbb{R}}\frac{dN_k^-(\lambda)}{\lambda-E}-e^{-i\theta(E)}=\theta'_{p,L}(E)f_k^-(E)(\cot u_L(E)-g_k^-(E)),\tag{5-20}$$

where g_k^- is as defined in (5-16).

Thus,

$$\left|\frac{\theta'_{p,L}(E)f_k^-(E)e^{-iu_L(E)}}{\sin u_L(E)}-\int_{\mathbb{R}}\frac{dN_k^-(\lambda)}{\lambda-E}-e^{-i\theta(E)}\right|\gtrsim|\cot u_L(E)-g_k^-(E)|$$

as, for η sufficiently small and $L\ge 1$, one has

$$0<\min_{\substack{\operatorname{Re}E\in I\\ \operatorname{Im}E\in[-\eta/L,0)}}|\theta'_{p,L}(E)f_k^-(E)|\le\max_{\substack{\operatorname{Re}E\in I\\ \operatorname{Im}E\in[-\eta/L,0)}}|\theta'_{p,L}(E)f_k^-(E)|<+\infty.$$

Now notice that by Corollary 4.4, for $E\in I$, one has

$$\operatorname{Im}\left(\int_{\mathbb{R}}\frac{dN_k^-(\lambda)}{\lambda-E}\right)=-\theta'_{p,L}(E)f_k^-(E)=-\frac{1}{\pi}n_k^-(E).\tag{5-21}$$

Thus, as $E\mapsto\operatorname{Im} e^{-i\theta(E)}$ is positive on I, the analytic function $E\mapsto g_k^-(E)$ has positive imaginary part larger than, say $2\tilde\eta$ on I; hence, it has imaginary part larger than, say, $\tilde\eta$ in some neighborhood of $I + \overline{D(0, \eta_0)}$ (for sufficiently small $\eta_0 > 0$). Let M be the maximum modulus of this function on $I + \overline{D(0, \eta_0)}$. Then, as $\max_{\operatorname{Re} E \in I,\, \operatorname{Im} E \in [-\eta_0/L, 0)} |\theta'_{p,L}(E)| \lesssim 1$, one has

$$\max_{\substack{\operatorname{Re} E \in I \\ \operatorname{Im} E \in [-\eta_0/L, 0) \\ |\cot(u_L(E))| < 2M}} |\operatorname{Im} \cot u_L(E)| \lesssim (M^2 + 1)\eta_0.$$

Possibly reducing η_0, this guarantees that, for $\operatorname{Re} E \in I$ and $\operatorname{Im} E \in [-\eta_0/L, 0)$, one has

$$|\cot u_L(E) - g_k^-(E)| \geq 2M - M \geq M \quad \text{or} \quad \operatorname{Im}(\cot u_L(E) - g_k^-(E)) \leq -\tilde{\eta} + \tfrac{1}{2}\tilde{\eta} = -\tfrac{1}{2}\tilde{\eta}.$$

This completes the proof of the first lower bound in (5-15) in Corollary 5.5.

To prove the second bound in (5-15), using (5-2) we compute

$$\begin{aligned} \frac{\det(\Gamma_L^{\text{eff}}(E) + e^{-i\theta(E)})}{n_k^-(E) n_0^+(E)} &= (\cot u_L(E) - g_k^-(E))(\cot u_L(E) - g_0^+(E)) - \frac{1}{\sin^2 u_L(E)} \\ &= -(g_0^+(E) + g_k^-(E)) \left(\cot u_L(E) - \frac{g_0^+(E) g_k^-(E) - 1}{g_0^+(E) + g_k^-(E)} \right), \end{aligned} \tag{5-22}$$

where g_k^- and g_0^+ are defined by (5-16) and (5-17).

Using Proposition 5.8, one then concludes the nonvanishing of $E \mapsto \det(\Gamma_L^{\text{eff}}(E) + e^{-i\theta(E)})$ in the complex rectangle $\{\operatorname{Re} E \in I,\ \operatorname{Im} E \in [-\eta_0/L, 0)\}$ (for η_0 sufficiently small) in the same way as above. This completes the proof of Corollary 5.5.

5B4. *The proof of Theorem 1.7.* To solve (2-4) and (2-8), by Theorem 5.1, we first solve the equations

$$\frac{\theta'_{p,L}(E) f_k^-(E) e^{-iu_L(E)}}{\sin u_L(E)} = \int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E} - e^{-i\theta(E)} \quad \text{and} \quad \det(\Gamma_L^{\text{eff}}(E) + e^{-i\theta(E)}) = 0 \tag{5-23}$$

in a rectangle $I + i[-\eta, -\tilde{\eta}/L]$. Indeed, in such a rectangle, by Theorem 5.1 equations (2-4) and (2-8) are equivalent to

$$\begin{aligned} \frac{\theta'_{p,L}(E) f_k^-(E) e^{-iu_L(E)}}{\sin u_L(E)} &= \int_{\mathbb{R}} \frac{dN_k^-(\lambda)}{\lambda - E} - e^{-i\theta(E)} + O(L^{-\infty}) \\ \text{and} \quad \det(\Gamma_L^{\text{eff}}(E) + e^{-i\theta(E)}) &= O(L^{-\infty}), \end{aligned} \tag{5-24}$$

respectively, where the terms $O(L^{-\infty})$ are analytic in a rectangle $\tilde{I} + i[-2\eta, -0)$ (where $I \subset \tilde{I}$) and the bound $O(L^{-\infty})$ holds in the supremum norm.

Thanks to (5-20) for $\bullet = \mathbb{N}$ and to (5-22) for $\bullet = \mathbb{Z}$, to solve the equations (5-23) it suffices to solve

$$\cot u_L(E) = c^{\bullet}(E), \tag{5-25}$$

where we recall $u_L(E) := (L - k)\theta_{p,L}(E)$, g_0^+ and g_k^- are as defined in (5-17) and (5-16), respectively, and, as in Section 1B3, we have set

- $c^{\mathbb{N}}(E) := g_k^-(E)$ in the case of the half-line,

- $c^{\mathbb{Z}}(E) := \dfrac{g_0^+(E)g_k^-(E)-1}{g_0^+(E)+g_k^-(E)}$ in the case of the line.

We want to solve (5-25) is a rectangle $I+i[-\varepsilon,0)$ for some ε small but fixed. Using Proposition 5.8, we pick ε so small that, in the rectangle $I+i[-\varepsilon,0]$, the only zeros of $c^\bullet - i$ are those on the real line and $\operatorname{Im} c^\bullet$ is positive in $I+i[-\varepsilon,0)$.

To solve (5-25), we change variables $u=(L-k)\theta_{p,L}(E)$, that is, we write

$$E=\theta_{p,L}^{-1}\left(\frac{u}{L-k}\right).$$

As, for L_0 sufficiently large, $\inf_{L\geq L_0,\, E\in I+i[-\varepsilon,0)} \operatorname{Re}\theta'_{p,L}(E) > c > 0$, at the cost of possibly reducing ε this real analytic change of variables maps $I+[-\varepsilon,\varepsilon]+i[-\varepsilon,0)$ into, say, D_L such that $I_L+i[-\eta(L-k),0]\subset D_L$ (for some $\eta>0$), where $I_L=(L-k)\theta_{p,L}\big(I+\big[-\tfrac{1}{2}\varepsilon,\tfrac{1}{2}\varepsilon\big]\big)$; the inverse change of variable maps $I_L+i[-\eta(L-k),0]$ into some domain, say $\widetilde{D}_L$, such that $I+[-\varepsilon',\varepsilon']+i[-\varepsilon',0]\subset\widetilde{D}_L$ (for some $0<\varepsilon'<\varepsilon$). Now, to find all the solutions to (5-25) in $I+i[-\varepsilon',0)$, we first solve the following equation in $I_L+i[-\eta(L-k),0]$:

$$\cot u = c^\bullet\circ\theta_{p,L}^{-1}\left(\frac{u}{L-k}\right) \tag{5-26}$$

As $u\mapsto\cot u$ is π-periodic, we split $I_L+i[-\eta(L-k),0]$ into vertical strips of the type

$$l\pi+[0,\pi]+i[-\eta(L-k),0],\quad l_-\leq l\leq l_+,\ (l_-,l_+)\in\mathbb{Z}^2.$$

Without loss of generality, we may assume that $I_L=[l_-,l_+]\pi$. To solve (5-26) on the rectangle $l\pi+[0,\pi]+i[-\eta(L-k),0]$, we shift u by $l\pi$ and solve the following equation on $[0,\pi]+i[-\eta(L-k),0]$:

$$\cot u = c^\bullet_{l,L}(u),\quad \text{where}\quad c^\bullet_{l,L}(\cdot):=c^\bullet\circ\theta_{p,L}^{-1}\left(\frac{\cdot+l\pi}{L-k}\right). \tag{5-27}$$

In proving Theorem 1.5, we have already shown that, for some $\tilde\eta>0$ (independent of L sufficiently large and $l_-\leq l\leq l_+$), (5-27) does not have a solution in $[0,\pi]+i[-\tilde\eta,0]$. The cotangent is an analytic one-to-one mapping from $[0,\pi)+i(-\infty,0]$ to $\mathbb{C}^+\setminus\{i\}$. Thus, for L sufficiently large and $\tilde\eta$ sufficiently small, the cotangent defines a one-to-one mapping from $[0,\pi)+i[-\eta(L-k),-\tilde\eta]$ onto $T_L=\overline{D(z_+,r_+)\setminus D(z_-,r_-)}$, analytic in the interior of $[0,\pi)+i[-\eta(L-k),-\tilde\eta]$ and continuous up to the boundary, where we have defined

$$z_+=i\frac{e^{4\eta(L-k)}+1}{e^{4\eta(L-k)}-1},\quad z_-=i\frac{e^{4\tilde\eta}-1}{e^{4\tilde\eta}-1},\quad r_+=\frac{2e^{2\tilde\eta}}{e^{4\tilde\eta}-1},\quad r_-=\frac{2e^{2\eta(L-k)}}{e^{4\eta(L-k)}-1}.$$

Moreover, the boundaries $\{0\}+i[-\eta(L-k),-\tilde\eta]$ and $\{\pi\}+i[-\eta(L-k),-\tilde\eta]$ are mapped onto the interval $[z_-+ir_-,z_++ir_+]$.

Let $\widetilde{Z}^\bullet$ denote the finite set of zeros of $E\mapsto c^\bullet(E)-i$ in I. Then, by a Taylor expansion near the zeros of $c-i$, we know that, for η sufficiently small, there exist $\varepsilon_0>0$ and $\tilde k\geq 1$ such that, for L sufficiently large:

- For $\varepsilon \in (0, \varepsilon_0)$, there exists $0 < \eta_-$ such that, for $l_- \le l \le l_+$, if one has
$$\left|\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right) - \widetilde{E}\right| \ge \varepsilon \quad \text{for all } \widetilde{E} \in \widetilde{Z}^\bullet,$$
then one has $\eta_- \le |\operatorname{Im} c_{l,L}^\bullet(u) - 1|$ for all $u \in [0, \pi] + i[-\eta(L-k), 0]$.
- For $u \in [0, \pi] + i[-\eta(L-k), 0]$ and $\widetilde{E}$ the point in $\widetilde{Z}^\bullet$ closest to $\theta_{p,L}^{-1}(l\pi/(L-k))$, one has
$$\varepsilon_0 \le (1 - \operatorname{Im} c_{l,L}^\bullet(u)) \cdot \left[\left|\theta_{p,L}^{-1}\left(\frac{\operatorname{Re} u + l\pi}{L-k}\right) - \widetilde{E}\right| + \frac{|\operatorname{Im} u|}{L-k}\right]^{-\tilde{k}} \le \frac{1}{\varepsilon_0}, \tag{5-28}$$
where $\tilde{k}$ is the order of $\widetilde{E}$ as a zero of $E \mapsto c^\bullet(E) - i$.

As a consequence of the above description of $c_{l,L}^\bullet$, we obtain:

Lemma 5.9. *There exists $\tilde{\eta}$ and η small such that, for L sufficiently large, for all $l_- \le l \le l_+$, $u \mapsto c_{l,L}^\bullet(u)$ maps the rectangle $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$ into a compact subset of $D(z_+, r_+) \setminus D(z_-, r_-)$ in such a way that*
$$\sup_{u \in \partial([0,\pi]+i[-\eta(L-k),-\tilde{\eta}])} |\cot u - c_{l,L}^\bullet(u)| \gtrsim \left(\left|\widetilde{E} - \theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right)\right| + \frac{\tilde{\eta}}{L-k}\right)^{\tilde{k}}, \tag{5-29}$$
where $\widetilde{E}$ is the root of $E \mapsto c^\bullet(E) - i$ closest to $\theta_{p,L}^{-1}(l\pi/(L-k))$ and $\tilde{k}$ is the order of this root.

Note that, under the assumptions of Lemma 5.9, (5-29) implies that
$$\sup_{u \in \partial([0,\pi]+i[-\eta(L-k),-\tilde{\eta}])} |\cot u - c_{l,L}^\bullet(u)| \gtrsim L^{-\tilde{k}}.$$
Thus we can define the analytic mapping $\cot^{-1} \circ c_{l,L}^\bullet$ on $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$; it maps the rectangle $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$ into a compact subset of $(0, \pi) + i(-\eta(L-k), -\tilde{\eta})$. Equation (5-27) on $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$ is, thus, equivalent to the fixed point equation on the same rectangle,
$$u = \cot^{-1} \circ c_{l,L}^\bullet(u) \tag{5-30}$$

We note that, for $\alpha \in (0, 1)$ and L sufficiently large, if for some $\widetilde{E} \in \widetilde{Z}^\bullet$ of multiplicity $\tilde{k}$ one has $|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| < L^{-\alpha}$, then (5-27) has no solution in $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$ outside of the set
$$R_{l,L} := [0, \pi] + i\left[-\eta(L-k), \frac{\alpha\tilde{k}}{4} \log\left[\left|\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right) - \widetilde{E}\right| + \frac{1}{L}\right]\right].$$

Indeed, for $u \in ([0, \pi] + i[-\eta(L-k), -\tilde{\eta}]) \setminus R_{l,L}$, by (5-28), that is, for
$$0 \le \operatorname{Re} u \le \pi \quad \text{and} \quad -\frac{\alpha\tilde{k}}{4} \log L \le \frac{\alpha\tilde{k}}{4} \log\left[\left|\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right) - \widetilde{E}\right| + \frac{1}{L}\right] \le \operatorname{Im} u \le -\tilde{\eta},$$
one has $|c_{l,L}^\bullet(u) - i| \lesssim L^{-\alpha\tilde{k}}$ and $|\cot u - i| \gtrsim L^{-\alpha\tilde{k}/2}$.

So if for some $\widetilde{E} \in \widetilde{Z}^{\bullet}$ one has $|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| < L^{-\alpha}$, it suffices to solve (5-30) on $R_{l,L}$. We compute the derivative of $c_{l,L}^{\bullet}$ in the interior of $R_{l,L}$:

$$\begin{aligned}\frac{d}{du}(\cot^{-1} \circ c_{l,L}^{\bullet})(u) &= -\frac{1}{L-k}\frac{c' \circ \theta_{p,L}^{-1}((u+l\pi)/(L-k))}{1+(c_{l,L}^{\bullet}(u))^2} \cdot \frac{1}{\theta'_{p,L}\big(\theta_{p,L}^{-1}((u+l\pi)/(L-k))\big)} \\ &= \frac{1}{L-k}\frac{c' \circ \theta_{p,L}^{-1}((u+l\pi)/(L-k))}{c_{l,L}^{\bullet}(u)-i} \cdot \frac{1}{c_{l,L}^{\bullet}(u)+i} \cdot \frac{1}{\theta'_{p,L}\big(\theta_{p,L}^{-1}((u+l\pi)/(L-k))\big)}.\end{aligned}$$

Thus, fixing $\alpha \in (0, 1)$:

- If l is such that for some $\widetilde{E} \in \widetilde{Z}^{\bullet}$ one has $|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| < L^{-\alpha}$, then for $u \in R_{l,L}$ we estimate

$$\begin{aligned}&\left|\frac{d}{du}(\cot^{-1} \circ c_{l,L}^{\bullet})(u)\right| \\ &\qquad \lesssim \frac{1}{L-k}\left[\left|\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right) - \widetilde{E}\right| + \frac{|\operatorname{Im} u|}{L-k}\right]^{-1} \\ &\qquad \lesssim \frac{1}{(L-k)|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| + \big|\log\big[|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| + \tilde{\eta}/(L-k)\big]\big|} \\ &\qquad \lesssim \frac{1}{\log L}.\end{aligned} \tag{5-31}$$

- If l is such that for all $\widetilde{E} \in \widetilde{Z}^{\bullet}$ one has $|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}| \geq L^{-\alpha}$, for $u \in [0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$ we estimate

$$\begin{aligned}\left|\frac{d}{du}(\cot^{-1} \circ c_{l,L}^{\bullet})(u)\right| &\lesssim \frac{1}{L-k}\left[\left|\theta_{p,L}^{-1}\left(\frac{l\pi}{L-k}\right) - \widetilde{E}\right| + \frac{|\operatorname{Im} u|}{L-k}\right]^{-1} \\ &\lesssim \frac{1}{(L-k)|\theta_{p,L}^{-1}(l\pi/(L-k)) - \widetilde{E}|} \lesssim \frac{1}{L^{1-\alpha}}.\end{aligned} \tag{5-32}$$

Hence, for L sufficiently large, $\cot^{-1} \circ c_{l,L}^{\bullet}$ is a contraction on $R_{l,L}$. Equation (5-30) thus admits a unique solution, say $\tilde{u}_{l,L}^{\bullet}$, in the rectangle $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$. This solution is a simple root of $u \mapsto u - \cot^{-1} \circ c_{l,L}^{\bullet}(u)$. Hence, $\tilde{u}_{l,L}^{\bullet}$ is the only solution to (5-27) in $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$.

By (5-24), for L sufficiently large and $l_- \leq l \leq l_+$, both the equations

$$\begin{aligned}S_L \circ \theta_{p,L}^{-1}\left(\frac{u+l\pi}{L-k}\right) + e^{-i\theta(\theta_{p,L}^{-1}((u+l\pi)/(L-k)))} &= 0, \\ \det\left(\Gamma_L \circ \theta_{p,L}^{-1}\left(\frac{u+l\pi}{L-k}\right) + e^{-i\theta(\theta_{p,L}^{-1}((u+l\pi)/(L-k)))}\right) &= 0,\end{aligned} \tag{5-33}$$

can be rewritten as

$$u = \cot^{-1}(c_{l,L}^{\bullet}(u) + O(L^{-\infty})) = \cot^{-1} \circ c_{l,L}^{\bullet}(u) + O(L^{-\infty}) \tag{5-34}$$

in $[0, \pi] + i[-\eta(L-k), -\tilde{\eta}]$.

Thus each of the equations in (5-33) admits a single solution in $[0,\pi]+i[-\eta(L-k),-\tilde\eta]$ and this root is simple; moreover, this solution, say $u_{l,L}$, satisfies $|u^\bullet_{l,L}-\tilde u^\bullet_{l,L}|=O(L^{-\infty})$; indeed, the bounds (5-31) and (5-32) guarantee that one can apply Rouché's theorem on the disk $D(\tilde u^\bullet_{l,L},L^{-k})$ for any $k\ge 0$.

Thus, we have proved:

Lemma 5.10. *Pick I as above. Then there exists $\eta>0$ such that, for L sufficiently large with $L=Np+k$, the resonances in $I+i[-\eta,0]$ are the energies $(z^\bullet_l)_{l_-\le l\le l_+}$ defined by*

$$z^\bullet_l=\theta^{-1}_{p,L}\left(\frac{u^\bullet_{l,L}+l\pi}{L-k}\right), \tag{5-35}$$

belonging to $I+i[-\eta,0]$.

Let us complete the proof of Theorem 1.14, that is, prove that, for η sufficiently small and L sufficiently large such that $L\equiv k \bmod p$, $z^\bullet_l$ is the unique resonance in $\left[\frac12\operatorname{Re}(\tilde z^\bullet_l+\tilde z^\bullet_{l-1}),\frac12\operatorname{Re}(\tilde z^\bullet_l+\tilde z^\bullet_{l+1})\right]+i[-\eta,0]$; recall that $\tilde z^\bullet_l$ is defined in (1-9).

We first note that the Taylor expansion of $\theta^{-1}_{p,L}$, (4-1) and the quantization condition (4-3) imply that

$$z^\bullet_l=\lambda_l+\frac{1}{\pi n(\lambda_l)L}u^\bullet_{l,L}+O\left(\left(\frac{\log L}{L}\right)^2\right)$$

as $\operatorname{Re}u_{l,L}\in[0,\pi)$ and $-\log L\lesssim\operatorname{Im}u_{l,L}\lesssim-1$.

Moreover, as

$$c^\bullet_{l,L}(u)=c^\bullet\left[\lambda_l+\frac{u}{\pi n(\lambda_l)L}+O\left(\frac{u^2}{L^2}\right)\right],$$

using (1-9) and (5-35) we compute

$$z^\bullet_l-\tilde z^\bullet_l=\frac{1}{\pi n(\lambda_l)L}\left(u^\bullet_{l,L}-\cot^{-1}\circ c^\bullet\left[\lambda_l+\frac{1}{\pi n(\lambda_l)L}\cot^{-1}\circ c^\bullet\left(\lambda_l-i\frac{\log L}{L}\right)\right]\right)+O\left(\left(\frac{\log L}{L}\right)^2\right).$$

Thus, one has

$$z^\bullet_l-\tilde z^\bullet_l=\frac{1}{\pi n(\lambda_l)L}\big(u^\bullet_{l,L}-\cot^{-1}\circ c^\bullet_{l,L}\big[\cot^{-1}\circ c^\bullet_{l,L}(-i\pi n(\lambda_l)\log L)\big]\big)+O\left(\left(\frac{\log L}{L}\right)^2\right).$$

As $u_{l,L}$ solves (5-34), sing (5-31) and (5-32) we thus obtain that

$$\begin{aligned}|z^\bullet_l-\tilde z^\bullet_l|&\lesssim\frac{1}{L\log L}\big|u^\bullet_{l,L}-\cot^{-1}\circ c^\bullet_{l,L}(-i\pi n(\lambda_l)\log L)\big|+\left(\frac{\log L}{L}\right)^2\\&\lesssim\frac{|u^\bullet_{l,L}|+\log L}{L\log^2 L}+\left(\frac{\log L}{L}\right)^2\lesssim\frac{1}{L\log L},\end{aligned}$$

using again $\operatorname{Re}u_{l,L}\in[0,\pi)$ and $-\log L\lesssim\operatorname{Im}u_{l,L}\lesssim-1$.

Taking into account (1-10), this completes the proof of Theorem 1.7.

5B5. *The proofs of Propositions 1.8 and 1.9.* Proposition 1.9 is an immediate consequence of Theorem 1.7, the definition (1-9) of $\tilde{z}_l^\bullet$ and the standard asymptotics of cot near $-i\infty$, i.e., $\cot z = i + 2ie^{-2iz} + O(e^{-4iz})$.

To prove Proposition 1.8, it suffices to notice that, under the assumptions of Proposition 1.8, the bound (5-32) on the derivative of $\cot^{-1} \circ c_{l,L}^\bullet$ on the rectangle $R_{l,L}$ becomes

$$\left| \frac{d}{du} (\cot^{-1} \circ c_{l,L}^\bullet)(u) \right| \lesssim \frac{1}{L}.$$

Thus, as a solution to (5-30), $u_{l,L}^\bullet$ admits an asymptotic expansion in inverse powers of L. Plugging this into (5-35) yields the asymptotic expansion for the resonance. Then (1-11) follows from the computation of the first terms.

5B6. *The proof of Theorem 1.10.* Theorem 1.10 is an immediate consequence of Theorem 5.2, the fact that the functions are analytic in the lower complex half-plane and have only finitely many zeros there, and the argument principle.

5C. ***The half-line periodic perturbation: the proof of Theorem 1.11.*** Using the same notations as above, we can write

$$H^\infty = \begin{pmatrix} H_{-1}^- & |\delta_{-1}\rangle\langle\delta_0| \\ |\delta_0\rangle\langle\delta_{-1}| & -\Delta_0^+ \end{pmatrix},$$

where $-\Delta_0^+$ is the Dirichlet Laplacian on $\ell^2(\mathbb{N})$.

Define the operators

$$\Gamma(E) := H_{-1}^- - E - \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle |\delta_{-1}\rangle\langle\delta_{-1}|,$$
$$\widetilde{\Gamma}(E) := -\Delta_0^+ - E - \langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle |\delta_0\rangle\langle\delta_0|.$$

For $\operatorname{Im} E \neq 0$, $\langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle$ and $\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle$ have a nonvanishing imaginary part of the same sign; hence, the complex number

$$\big(\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle\big)^{-1} - \langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle$$

does not vanish. Thus, by rank-one perturbation theory, (see, e.g., [Simon 2005]), we know that $\Gamma(E)$ and $\widetilde{\Gamma}(E)$ are invertible and their inverses are given by

$$\Gamma^{-1}(E) := (H_{-1}^- - E)^{-1} + \frac{\big|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle\langle\delta_{-1}|(H_{-1}^- - E)^{-1}\big|}{\big(\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle\big)^{-1} - \langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle} \tag{5-36}$$

and

$$\widetilde{\Gamma}^{-1}(E) := (-\Delta_0^+ - E)^{-1} + \frac{\big|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle\langle\delta_0|(-\Delta_0^+ - E)^{-1}\big|}{\big(\langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle\big)^{-1} - \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle}. \tag{5-37}$$

Thus, for $\operatorname{Im} E \neq 0$, using Schur's complement formula we compute

$$(H^\infty - E)^{-1} = \begin{pmatrix} \Gamma(E)^{-1} & \gamma(E) \\ \gamma^*(\bar{E}) & \widetilde{\Gamma}(E)^{-1} \end{pmatrix}, \tag{5-38}$$

where $\gamma^*(\bar{E})$ is the adjoint of $\gamma(\bar{E})$ and

$$\gamma(E) := -|\Gamma(E)^{-1}|\delta_{-1}\rangle\langle\delta_0|(-\Delta_0^+ - E)^{-1}|.$$

Now, when coming from $\operatorname{Im} E > 0$ and passing through $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$, the complex numbers

$$\langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle \quad \text{and} \quad \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle$$

keep imaginary parts of the same positive sign; thus, the two operator-valued functions $E \mapsto \Gamma^{-1}(E)$ and $E \mapsto (H^\infty - E)^{-1}$ can be analytically continued through $(-2, 2) \cap \Sigma_{\mathbb{Z}}^\circ$ from the upper to the lower complex half-plane (as operators from $\ell^2_{\text{comp}}(\mathbb{N})$ to $\ell^2_{\text{loc}}(\mathbb{N})$ and from $\ell^2_{\text{comp}}(\mathbb{Z})$ to $\ell^2_{\text{loc}}(\mathbb{Z})$, respectively).

When coming from the upper half-plane and passing through $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ and $\Sigma_{\mathbb{Z}}^\circ \setminus [-2, 2]$, (5-38) also provides an analytic continuation of $(H^\infty - E)^{-1}$. Definition (5-36) and formula (5-38) immediately show that the poles of these continuations only occur at the zeros of the function

$$E \mapsto 1 - \langle\delta_{-1}|(H_{-1}^- - E)^{-1}|\delta_{-1}\rangle\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle = 1 - e^{i\theta(E)} \int_{\mathbb{R}} \frac{dN_{p-1}^-(\lambda)}{\lambda - E}$$

when continued from the upper half-plane through the sets $(-2, 2) \setminus \Sigma_{\mathbb{Z}}$ and $\Sigma_Z^\circ \setminus [-2, 2]$ (these sets are finite unions of open intervals).

This completes the proof of Theorem 1.11.

6. Resonances in the random case

As for the periodic potential, for the random potential we start with a description of the function $E \mapsto \Gamma_L(E)$ (see (2-9)), that is, with a description of the spectral data for the Dirichlet operator $H_{\omega,L}$.

6A. ***The matrix Γ_L in the random case.*** We recall a number of results on the Dirichlet eigenvalues of $H_{\omega,L}$ that will be used in our analysis.

It is well known that, under our assumptions, in dimension one the whole spectrum of H_ω is in the localization region (see, e.g., [Kunz and Souillard 1980; Cycon et al. 1987; Carmona and Lacroix 1990]), that is:

Theorem 6.1. *There exists $\rho > 0$ and $\alpha \in (0, 1)$ such that one has*

$$\sup_{\substack{L \in \mathbb{N} \cup \{+\infty\} \\ y \in [\![0, L]\!] \\ \operatorname{Im} E \neq 0}} \mathbb{E}\left\{ \sum_{x \in [\![0, L]\!]} e^{\rho|x-y|} |\langle\delta_x, (H_{\omega,L} - E)^{-1}\delta_y\rangle|^\alpha \right\} < \infty \tag{6-1}$$

and

$$\sup_{\substack{L \in \mathbb{N} \cup \{+\infty\} \\ y \in [\![0, L]\!]}} \mathbb{E}\left\{ \sum_{x \in [\![0, L]\!]} e^{\rho|x-y|} \sup_{\substack{\operatorname{supp} f \subset \mathbb{R} \\ |f| \le 1}} |\langle\delta_x, f(H_{\omega,L})\delta_y\rangle| \right\} < \infty, \tag{6-2}$$

where $H_{\omega,+\infty} := H_\omega^{\mathbb{N}}$ and $[\![0, +\infty]\!] = \mathbb{N}$. The supremum is taken over the functions f that are Borelian and compactly supported.

As a consequence, one can define localization centers, e.g., by means of the following results:

Lemma 6.2 [Germinet and Klopp 2014]. *Fix $(l_L)_L$ a sequence of scales, i.e., $l_L \to +\infty$ as $L \to +\infty$. There exists $\rho > 0$ such that, for L sufficiently large, with probability larger than $1 - e^{-\ell_L}$, if*

(1) *$\varphi_{j,\omega}$ is a normalized eigenvector of $H_{\omega,L}$ associated to $E_{j,\omega}$ in Σ,*

(2) *$x_j(\omega) \in [\![0, L]\!]$ is a maximum of $x \mapsto |\varphi_{j,\omega}(x)|$ in $[\![0, L]\!]$,*

then for $x \in [\![0, L]\!]$ one has

$$|\varphi_{j,\omega}(x)| \le \sqrt{L} e^{2\ell_L} e^{-\rho|x - x_j(\omega)|}. \tag{6-3}$$

Note that Lemma 6.2 is of interest only if $\ell_L \lesssim L$; otherwise (6-3) is obvious. This result can, for example, be applied for the scales $l_L = 2 \log L$. In this case, the probability estimate of the bad sets (i.e., when the conclusions of Lemma 6.3 does not hold) is summable. The point $x_j(\omega)$ is a localization center for $E_{j,\omega}$ or $\varphi_{j,\omega}$. It is not defined uniquely, but, one easily shows that there exists $C > 0$ such that for any two localization centers, say x and x', one has $|x - x'| \le C \log L$ (see [Germinet and Klopp 2014]). For concreteness, we set the localization center associated to the eigenvalue $E_{j,\omega}$ to be the leftmost maximum of $x \mapsto \|\varphi_{j,\omega}\|_x$.

We show:

Lemma 6.3. *For any $p > 0$, there exist $C > 0$ and $L_0 > 0$ (depending on α and p) such that, for $L \ge L_0$, for any sequence satisfying* (1-22), *with probability at least $1 - L^{-p}$ there exist at most $C\ell_L$ eigenvalues having a localization center in $[\![0, \ell_L]\!] \cup [\![L - \ell_L, L]\!]$.*

We will now use the fact that we are dealing with one-dimensional systems to improve upon the estimate (6-3). We prove:

Theorem 6.4. *For any $\delta > 0$ and $p \ge 0$, there exist $C > 0$ and $L_0 > 0$ (depending on p and δ) such that, for $L \ge L_0$, with probability at least $1 - L^{-p}$ if $E_{j,\omega}$ is an eigenvalue in Σ associated to the eigenfunction $\varphi_{j,\omega}$ and the localization center $x_{j,\omega}$ then:*

- *If $x_{j,\omega} \in [\![0, L - C \log L]\!]$, one has*

$$-\rho(E_{j,\omega}) - \delta \le \frac{\log |\varphi_{j,\omega}(L)|}{L - x_{j,\omega}} \le -\rho(E_{j,\omega}) + \delta. \tag{6-4}$$

- *If $x_{j,\omega} \in [\![C \log L, L]\!]$, one has*

$$-\rho(E_{j,\omega}) - \delta \le \frac{\log |\varphi_{j,\omega}(0)|}{x_{j,\omega}} \le -\rho(E_{j,\omega}) + \delta. \tag{6-5}$$

To analyze the resonances of $H^{\mathbb{N}}_{\omega,L}$ (resp. $H^{\mathbb{Z}}_{\omega,L}$), we shall use (6-4) (resp. (6-4) and (6-5)).

We now use these estimates as the starting point of a short digression from the main theme of this paper. Let us first state a corollary to Theorem 6.4; we prove:

Theorem 6.5. *For any $\delta > 0$ and $p \ge 0$, for L sufficiently large (depending on p and δ), with probability at least $1 - L^{-p}$, if $E_{j,\omega}$ is an eigenvalue in Σ associated to the eigenfunction $\varphi_{j,\omega}$ and the localization* *center $x_{j,\omega}$ then, for $|x-x_{j,\omega}|\geq\delta L$ and $1\leq x\leq L$, one has*

$$-\rho(E_{j,\omega})-\delta\leq\frac{\log\big(|\varphi_{j,\omega}(x)|+|\varphi_{j,\omega}(x-1)|\big)}{|x-x_{j,\omega}|}\leq-\rho(E_{j,\omega})+\delta. \tag{6-6}$$

Compare (6-6) to (6-3). There are two improvements. First, the unknown rate of decay ρ is replaced by the Lyapunov exponent $\rho(E_{j,\omega})$, which was expected to be the correct decay rate. Indeed, for the one-dimensional discrete Anderson model on the half-axis, it is well known (see, e.g., [Bougerol and Lacroix 1985; Carmona and Lacroix 1990; Pastur and Figotin 1992]) that, ω-almost surely, the spectrum is localized and the eigenfunctions decay exponentially at infinity at a rate given by the Lyapunov exponent. In Theorem 6.5, we state that, with good probability, this is true for finite volume restrictions.

Second, in (6-6), we get both an upper and lower bound on the eigenfunction. This is more precise than (6-3).

To our knowledge, such a result was not known until the present paper. The strategy that we use to prove this result can be applied in a more general one-dimensional setting to obtain analogues of (6-6) (see [Klopp $\geq$ 2016]).

We complement this with the much simpler:

Lemma 6.6. *For any $C>0$ and $p\geq 0$, there exists $K>0$ and $L_0>0$ (depending on p and C) such that, for $L\geq L_0$, with probability at least $1-L^{-p}$ if $E_{j,\omega}$ is an eigenvalue in Σ associated to the eigenfunction $\varphi_{j,\omega}$ and the localization center $x_{j,\omega}$ then:*

- *If $x_{j,\omega}\in[\![L-C\log L, L]\!]$, one has $L^{-K}\leq|\varphi_{j,\omega}(L)|$.*
- *If $x_{j,\omega}\in[\![0, C\log L]\!]$, one has $L^{-K}\leq|\varphi_{j,\omega}(0)|$.*

The proof of this result is obvious and only uses the fact that the matrices in the cocycle defining the operator (see Section 6C) are bounded, that is, equivalently, that the solutions to the Schrödinger equation grow at most exponentially at a rate controlled by the potential.

Let us return to the resonances in the random case and the description of the function S_L. Recall that in (2-4) the values $(\lambda_j)_j$ are the eigenvalues $(E_{j,\omega})_{0\leq j\leq L}$ of $H_{\omega,L}$ and the coefficients $(a^\bullet_j)_j$ are defined in Theorem 2.2 and by (2-13). Thus, Theorem 6.4 describes the coefficients $(a^\bullet_j)_j$ coming into S_L and Γ_L (see (2-4) and (2-8)). Let us now state a few consequences of Theorem 6.4.

Fix a compact interval I in Σ, the almost sure spectrum of H_ω. For $\bullet\in\{\mathbb{N},\mathbb{Z}\}$, define

$$d^\bullet_{j,\omega}=\begin{cases}L-x_{j,\omega} & \text{for } \bullet=\mathbb{N},\\ \min(x_{j,\omega}, L-x_{j,\omega}) & \text{for } \bullet=\mathbb{Z}.\end{cases} \tag{6-7}$$

Taking $p>2$ in Theorem 6.4 and using a Borel–Cantelli argument, we obtain:

ω-almost surely, for $\delta>0$ and L sufficiently large,

$$\text{if } \lambda_j=E_{j,\omega}\in I \text{ and } d^\bullet_{j,\omega}\geq C\log L \text{ then } -2\rho(\lambda_j)-\delta\leq\frac{\log a^\bullet_j}{d^\bullet_{j,\omega}}\leq-2\rho(\lambda_j)+\delta. \tag{6-8}$$

This and the continuity of the Lyapunov exponent (see, e.g., [Bougerol and Lacroix 1985; Carmona and Lacroix 1990; Pastur and Figotin 1992]) guarantees that

$$\omega\text{-almost surely, for any } \delta > 0 \text{ and } L \text{ large, one has } -2\eta_\bullet \sup_{E\in I} \rho(E)(1+\delta)L \le \inf_{\lambda_j\in I} \log a_j^\bullet, \tag{6-9}$$

where $\eta_\bullet$ is as defined in Theorem 1.13.

To use the analysis performed in Section 3, we also need a description for the $(\lambda_j)_j$, i.e., the Dirichlet eigenvalues of $H_{\omega,L}$. To this end, we will use the results of [Germinet and Klopp 2014; Klopp 2011; 2013] (see also [Germinet and Klopp 2011]).

We first recall the Minami estimate satisfied by $H_{\omega,L}$ (see, e.g., [Combes et al. 2009] and references therein): there exists $C > 0$ such that, for $I \subset \mathbb{R}$, one has

$$\mathbb{P}\big(\mathrm{tr}(\mathbf{1}_I(H_{\omega,L})) \ge 2\big) \le \mathbb{E}\big(\mathrm{tr}(\mathbf{1}_I(H_{\omega,L}))\big[\mathrm{tr}(\mathbf{1}_I(H_{\omega,L})) - 1\big]\big) \le C|I|^2(L+1)^2.$$

Here $\mathbf{1}_I(H)$ denotes the spectral projector for the selfadjoint operator H onto the energy interval I.

By a simple covering argument, this entails the estimate

$$\mathbb{P}(|\lambda_i - \lambda_j| \le L^{-q} \text{ for some } i \ne j) \le CL^{-q+2}.$$

Thus, for $q > 3$, a Borel–Cantelli argument yields that

$$\omega\text{-almost surely, for } L \text{ sufficiently large, } \min_{i\ne j} |\lambda_i - \lambda_j| \ge L^{-q}. \tag{6-10}$$

6B. *The proofs of the main results in the random case.* We are now going to prove the results stated in Section 1C.

6B1. *The proof of Theorem 1.13.* As for Theorem 1.5, this result follows from Theorem 3.1. Point (1) is proved exactly as Theorem 1.5(1). Point (2) follows immediately from Theorem 3.1 and (6-9). This completes the proof of Theorem 1.13.

6B2. *The proof of Theorem 1.14.* Recall that $\kappa \in (0,1)$. To prove (1) we proceed as follows. The standard result guaranteeing the existence of the density of states N (see, e.g., [Bougerol and Lacroix 1985; Carmona and Lacroix 1990; Pastur and Figotin 1992]) implies that, ω-almost surely, one has

$$\frac{\#\{\lambda_j \in I\}}{L+1} \to \int_I dN(E). \tag{6-11}$$

This, in particular, shows that if $I \subset \Sigma^\circ$ is a compact interval then, ω-almost surely, for L sufficiently large I is covered by intervals of the form $[\lambda_j, \lambda_{j+1}]$ and their number is of size $\asymp L$ (actually this holds for $\lambda_j \in I + [-\varepsilon, \varepsilon]$ if $\varepsilon > 0$ is chosen small enough). Moreover, the estimate (6-10) guarantees that $d_j \ge L^{-q}$ (for any $q > 3$ fixed) for all $\lambda_j \in I$. Thus, Theorems 3.1, 3.2 and 3.3 and the estimate (6-8) guarantee that, ω-almost surely, all the resonances in the strip $I - i[e^{-L^\kappa}, 0)$ are described by Theorem 3.3. Indeed, for such a resonance the imaginary part must be larger than $-e^{-L^\kappa}$; thus, by Theorem 3.1, for every rectangle $\big[\frac{1}{2}(\lambda_j+\lambda_{j-1}), \frac{1}{2}(\lambda_j+\lambda_{j+1})\big] - i[e^{-L^\kappa}, 0)$ containing a resonance, one has $a_j^\bullet \lesssim e^{-L^\kappa} L^{2q}$ Thus, $a_j^\bullet \ll d_j^2$ and one can apply Theorem 3.3 to compute the resonance.

Let us count the number of those resonances. To this end, let $\ell_L = \tau L^\kappa$, where τ is to be chosen. By (6-8) and (6-10), ω-almost surely one has $a^\bullet_j \ll d_j^2$ for all j such that $\lambda_j \in I$ as long as the Dirichlet eigenvalue λ_j is associated to a localization center in $[\![0, L-\ell_L]\!]$ (actually this holds for $\lambda_j \in I+[-\varepsilon, \varepsilon]$ if $\varepsilon > 0$ is chosen small enough); thus, we can apply Theorems 3.3 and 3.2 to each of the $(\lambda_j)_j$ that are associated to a localization center in $[\![0, L-\ell_L]\!]$. By (3-19), each of these eigenvalues gives rise to a single simple resonance, the imaginary part of which is of size $\asymp a^\bullet_j$; they lie above the line $\{\operatorname{Im} z \geq e^{-\rho\ell_L} = e^{-L^\kappa}\}$ for $\tau\rho = 1$. Actually, the estimate (6-10) guarantees that $d_j \geq L^{-q}$ (for any $q > 3$ fixed) and Theorem 3.2 shows that these resonances are the only ones above the line $\operatorname{Im} z \geq -L^{-q}$. Moreover, by Lemma 6.3, we know there at most $C\ell_L$ eigenvalues λ_j that do not have their localization center in $[\![0, L-\ell_L]\!]$. Thus we obtain, ω-almost surely,

$$\lim_{L\to+\infty} \frac{1}{L} \#\big\{z \text{ resonance of } H_{\omega,L} \text{ with } \operatorname{Re} z \in I,\ \operatorname{Im} z \geq -e^{-L^\kappa}\big\} = \int_I dN(E).$$

Point (2) is proved in the same way. Pick $\lambda \in (0, 1)$. In addition to what was used above, one uses the continuity of the density of states $E \mapsto n(E)$ and the Lyapunov exponent $E \mapsto \rho(E)$. Assume E is as in point (2). Then, ω-almost surely, the reasoning done above shows that, for any $\eta > 0$, there exists $\varepsilon_0 > 0$ such that, for $\varepsilon \in (0, \varepsilon_0)$ and $\delta \in (0, \delta_0)$, for L sufficiently large one has

$$\begin{aligned}\#\Big\{\lambda_l \text{ eigenvalue of } H^{\mathbb{N}}_{\omega,L} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1+\eta, 1-\eta] \text{ with } -e^{\eta_\bullet\rho(E)\delta L} \lesssim e^{2\eta_\bullet\rho(E)\lambda L} a^\bullet_l \lesssim -e^{-\eta_\bullet\rho(E)\delta L}\Big\}&\\ \leq \#\big\{z \text{ resonance of } H^\bullet_{\omega,L} \text{ in } R^\bullet(E, \lambda, L, \varepsilon, \delta)\big\}&\\ \leq \#\Big\{\lambda_l \text{ eigenvalue of } H^{\mathbb{N}}_{\omega,L} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1-\eta, 1+\eta]&\\ \text{with } -e^{\eta_\bullet\rho(E)\delta L} \lesssim e^{2\eta_\bullet\rho(E)\lambda L} a^\bullet_l \lesssim -e^{-\eta_\bullet\rho(E)\delta L}\Big\}&.\end{aligned}$$

Using Theorem 6.4 and the continuity of the Lyapunov exponent in conjunction with the definition of a_j (see (2-4) and (2-13)), we obtain that, ω-almost surely, for any $\eta > 0$ there exists $\varepsilon_0 > 0$ such that, for $\varepsilon \in (0, \varepsilon_0)$ and $\delta \in (0, \delta_0)$, for L sufficiently large one has

$$\begin{aligned}\#\Big\{\text{eigenvalue of } H^{\mathbb{N}}_{\omega,L} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1+\eta, 1-\eta] \text{with localization center in } I^\bullet(L, \delta, -\eta)\Big\}&\\ \leq \#\big\{z \text{ resonance of } H^\bullet_{\omega,L} \text{ in } R^\bullet(E, \lambda, L, \varepsilon, \delta)\big\}&\\ \leq \#\Big\{\text{eigenvalue of } H^{\mathbb{N}}_{\omega,L} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1-\eta, 1+\eta] \text{ with localization center in } I^\bullet(L, \delta, \eta)\Big\}&,\end{aligned}$$

where $I^{\mathbb{N}}(L, \lambda, \delta, \eta)$ is the interval — here $[r]$ denotes the integer part of $r \in \mathbb{R}$ —

$$I^{\mathbb{N}}(L, \lambda, \delta, \eta) = [L\lambda] + [\![-L\delta(1+\eta), L\delta(1+\eta)]\!]$$

and $I^{\mathbb{Z}}(L, \lambda, \delta, \eta)$ is the union of intervals

$$I^{\mathbb{Z}}(L, \lambda, \delta, \eta) = \big(\big[\tfrac{1}{2}L\lambda\big] + [\![-L\delta(1+\eta), L\delta(1+\eta)]\!]\big) \cup \big(\big[L\big(1-\tfrac{1}{2}\lambda\big)\big] + [\![-L\delta(1+\eta), L\delta(1+\eta)]\!]\big).$$

Now, using the exponential localization of the eigenfunctions, one has that, ω-almost surely, for any $\eta > 0$ there exists $\varepsilon_0 > 0$ such that, for $\varepsilon \in (0, \varepsilon_0)$ and $\delta \in (0, \delta_0)$, for L sufficiently large one has

$$\begin{aligned}\#\left\{\text{eigenvalue of } H^{\mathbb{N}}_{\omega,L,\lambda,\delta,-2\eta,\bullet} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1+2\eta,1-2\eta]\right\}\\ \le \#\{z \text{ resonance of } H^{\bullet}_{\omega,L} \text{ in } R^{\bullet}(E,\lambda,L,\varepsilon,\delta)\}\\ \le \#\left\{\text{eigenvalue of } H^{\mathbb{N}}_{\omega,L,\lambda,\delta,2\eta,\bullet} \text{ in } E+\frac{\varepsilon}{2n(E)}[-1-2\eta,1+2\eta]\right\},\end{aligned} \tag{6-12}$$

where $H^{\mathbb{N}}_{\omega,L,\lambda,\delta,\eta,\bullet} = (H^{\mathbb{N}}_{\omega,L})_{|I^{\bullet}(L,\lambda,\delta,\eta)}$ with Dirichlet boundary conditions at the edges of the interval $I^{\bullet}(L,\lambda,\delta,\eta)$.

This immediately yields point (2) for $\lambda \in (0, 1)$, using (6-11) for the operators $H^{\mathbb{N}}_{\omega,L,\lambda,\delta,\eta,\bullet}$. The case $\lambda = 1$ is dealt with in the same way.

As already said, point (3) is an “integral” version of point (2). Using the same ideas as above, partitioning $I = \bigcup_{p=0}^{P} I_p$ so that $|I_p| \sim \varepsilon$ centered in E_p, one proves

$$\begin{aligned}\sum_{p=0}^{P}\#\left\{\text{eigenvalue of } H^{-}_{\omega,p,L,\bullet} \text{ in } E_p+\frac{\varepsilon}{2n(E_p)}[-1+2\eta,1-2\eta]\right\}\\ \le \#\{z \text{ resonance of } H^{\bullet}_{\omega,L} \text{ in } I+[-e^{-L^{\kappa}},-e^{-cL}]\}\\ \le \sum_{p=0}^{P}\#\left\{\text{eigenvalue of } H^{+}_{\omega,p,L,\bullet} \text{ in } E_p+\frac{\varepsilon}{2n(E_p)}[-1-2\eta,1+2\eta]\right\},\end{aligned}$$

where

- $H^{-}_{\omega,p,L,\bullet}$ is the operator $H^{\mathbb{N}}_{\omega}$ restricted to
  - $[\![2L^{\kappa}, (\inf(c\rho^{-1}(E_p),1)-\eta)L]\!]$ if $\bullet = \mathbb{N}$,
  - $[\![2L^{\kappa}, (\inf(c\rho^{-1}(E_p),1)/2-\eta)L]\!] \cup [\![(1-\inf(c\rho^{-1}(E_p),1)/2+\eta)L, L-2L^{\kappa}]\!]$ if $\bullet = \mathbb{Z}$;
- $H^{+}_{\omega,p,L,\bullet}$ is the operator $H^{\mathbb{N}}_{\omega}$ restricted to
  - $[\![L^{\kappa}/2, (\inf(c\rho^{-1}(E_p),1)+\eta)L]\!]$ if $\bullet = \mathbb{N}$,
  - $[\![L^{\kappa}/2, (\inf(c\rho^{-1}(E_p),1)/2+\eta)L]\!] \cup [\![(1-\inf(c\rho^{-1}(E_p),1)/2-\eta)L, L-L^{\kappa}/2]\!]$ if $\bullet = \mathbb{Z}$.

In the computation above, we used the continuity of both the density of states $E \mapsto n(E)$ and the Lyapunov exponent $E \mapsto \rho(E)$. Thus, we obtain

$$\begin{aligned}\#\{z \text{ resonance of } H^{\bullet}_{\omega,L} \text{ in } I+(-\infty,e^{-cL}]\}\\ = L\left(\sum_{p=0}^{P}\inf(c\rho^{-1}(E_p),1)n(E_p)|I_p|+o(1)\right)+\#\{z \text{ resonance of } H^{\bullet}_{\omega,L} \text{ in } I+(-\infty,e^{-L^{\kappa}}]\}.\end{aligned}$$

The last term being controlled by Theorem 1.17, one obtains point (3) as the Riemann sum in the right-hand side above converges to the integral in the right-hand side of (1-18) as $\varepsilon \to 0$. This completes the proof of Theorem 1.14.

6B3. *The proof of Theorem 1.15.* The proof of Theorem 1.15 relies on [Germinet and Klopp 2014, Theorem 1.13], which describes the local distribution of the eigenvalues and localization centers $(E_{j,\omega}, x_{j,\omega})$; namely, one has

$$\lim_{L\to+\infty} \mathbb{P}\big(\big\{\omega \,\big|\, \#\{n \mid E_{j,\omega} \in E+L^{-1}I_n,\ x_{j,\omega} \in LC_n\} = k_n \ \text{ for } \ n=1,\dots,p\big\}\big) = \prod_{n=1}^{p} e^{-\tilde{\mu}_n} \frac{(\tilde{\mu}_n)^{k_n}}{k_n!}, \tag{6-13}$$

where $\tilde{\mu}_n := n(E)|I_n||C_n|$ for $1 \le n \le p$.

Recall that $(z_j^L(\omega))_j$ are the resonances of $H_{\omega,L}$. By the argument used in the proof of Theorem 1.14, we know that, ω-almost surely, all the resonances in $K_L := [E-\varepsilon, E+\varepsilon] + i[-e^{-L^\kappa}, 0]$ are constructed from the $(\lambda_j, a_j^\bullet)$ by formula (3-19). Thus, up to renumbering, the rescaled real and imaginary parts (see (1-19)) become

$$\begin{aligned} x_j &= (\operatorname{Re} z^\bullet_{l,L}(\omega) - E)L = (\lambda_j - E)L + O(La_j) = (E_{j,\omega} - E)L + O(Le^{-L^\kappa}),\\ y_j &= -\frac{1}{2L}\log|\operatorname{Im} z^\bullet_{l,L}(\omega)| = -\frac{\log a_j^\bullet}{2L} + O(1/L) = \rho(E)\frac{d^\bullet_{j,\omega}}{L} + o(1), \end{aligned}$$

where $\lambda_j = E_{j,\omega}$ and $d^\bullet_{j,\omega}$ is defined as in (6-7); here we used the continuity of $E \mapsto \rho(E)$.

On the other hand, for the resonances below the line in $\{\operatorname{Im} z = -e^{-L^\kappa}\}$, one has $y_j \lesssim L^{\kappa-1}$. So all these resonances are "pushed upwards" towards the upper half-plane. Hence, the statement of Theorem 1.15 is an immediate consequence of (6-13).

6B4. *The proof of Theorem 1.16.* Using the computations of the previous section, as $E \neq E'$, Theorem 1.16 is a direct consequence of [Klopp 2011, Theorem 1.2] (see also [Germinet and Klopp 2014, Theorem 1.11]).

6B5. *The proof of Theorem 1.17.* Consider equations (2-4) and (2-8). By Theorem 6.4 and Lemma 6.3, ω-almost surely, for L large the number of $(a_j^\bullet)_j$ larger than $e^{-10\ell_L}$ is bounded by $C\ell_L$. Solving (2-4) and (2-8) in the strip $\{\operatorname{Re} E \in I,\ \operatorname{Im} E < -e^{-\ell_L}\}$, we can write $S_L(E) = S_L^-(E) + S_L^+(E)$, where

$$S_L^-(E) := \sum_{a_j^{\mathbb{N}} \le e^{-10\ell_L}} \frac{a_j^{\mathbb{N}}}{\lambda_j - E} \quad \text{and} \quad S_L^+(E) := \sum_{a_j^{\mathbb{N}} > e^{-10\ell_L}} \frac{a_j^{\mathbb{N}}}{\lambda_j - E},$$

and similarly decompose $\Gamma_L(E) = \Gamma_L^-(E) + \Gamma_L^+(E)$. For L large, one then has

$$\sup_{\operatorname{Im} E < -e^{-\ell_L}} \|S_L^-(E)\| + \|\Gamma_L^-(E)\| \le e^{-8\ell_L}. \tag{6-14}$$

The count of the number of resonances given by the proof of Theorems 2.2 and 2.3 then shows that the equations (2-4) and (2-8), where S_L and Γ_L are respectively replaced by S_L^+ and Γ_L^+, have at most $C\ell_L$ solutions in the lower half-plane. We will call the equations where S_L and Γ_L are replaced by S_L^+ and Γ_L^+ the $+$-equations. The analogues of Theorems 3.1, 3.2 and 3.3 for the $+$-equations and Theorem 6.4 show that the only solutions to the $+$-equations in the strip $\{\operatorname{Re} E \in I,\ -e^{-4\ell_L/5} < \operatorname{Im} E < -e^{-3\ell_L/4}\}$ are given by formulas (3-19) and (3-20) for the eigenvalues of the Dirichlet problem associated to a localization center in $[\![L-2\ell_L, L-\frac{1}{2}\ell_L]\!]$ if $\bullet = \mathbb{N}$ and in $[\![\frac{1}{2}\ell_L, 2\ell_L]\!] \cup [\![L-2\ell_L, L-\frac{1}{2}\ell_L]\!]$ if $\bullet = \mathbb{Z}$. Thus, these

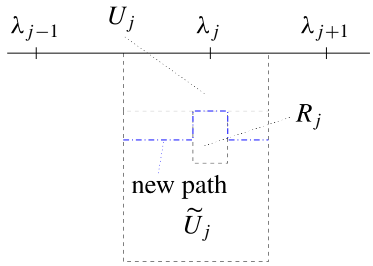


Figure 8. The new path (in blue).

zeros are simple and separated by a distance at least L^{-4} from each other (recall (6-10)). Moreover, we can cover the interval I by intervals of the type $\left[\frac{1}{2}(\lambda_j+\lambda_{j-1}), \frac{1}{2}(\lambda_j+\lambda_{j+1})\right]$, that is, one can write

$$I \subset \bigcup_{j=j^-}^{j^+} \left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right], \tag{6-15}$$

where $\lambda_{j^- -1} \notin I,\ \lambda_{1+j^+} \notin I,\ \lambda_{j^-} \in I$ and $\lambda_{j^+} \in I$.

Consider now the line $\{\operatorname{Im} E = -e^{-\ell_L}\}$ and its intersection with the vertical strip

$$\left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right] - i\mathbb{R}^+.$$

Three things may occur:

(1) $e^{-\ell_L} < a_j^\bullet d_j^2 |\sin\theta(\lambda_j)|/C$ (the constant C is defined in Theorem 3.1); then, on the interval

$$\left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right] - ie^{-\ell_L},$$

one has

$$|S_L^+(E) + e^{-i\theta(E)}| \gtrsim 1 \quad \text{and} \quad |\det(\Gamma_L^+(E) + e^{-i\theta(E)})| \gtrsim 1; \tag{6-16}$$

this follows from the proof of Theorem 3.1 (see in particular (3-5), (3-6), (3-7) and (3-8)) for some fixed $c > 0$; recall that, on the interval $I + ie^{-\ell_L}$, one has $|\sin\theta(E)| \gtrsim 1$.

(2) $e^{-\ell_L} > Ca_j^\bullet$ (the constant C is defined in Theorem 3.2); then, on the interval

$$\left[\tfrac{1}{2}(\lambda_j+\lambda_{j-1}), \tfrac{1}{2}(\lambda_j+\lambda_{j+1})\right] - ie^{-\ell_L},$$

one has again (6-16) for a possibly different constant; this follows from the proof of Theorem 3.2 (see in particular (3-15) and (3-16)).

(3) If we are neither in case (1) nor in case (2), then the line $\{\operatorname{Im} E = -e^{-\ell_L}\}$ may cross R_j (defined in Theorem 3.3; see also Figure 7); we change the contour $\{\operatorname{Im} E = -e^{-\ell_L}\}$ so as to enter $\widetilde{U}_j$ until we reach the boundary of R_j and then follow this boundary, getting closer to the real axis, turning around R_j and finally reaching the line $\{\operatorname{Im} E = -e^{-\ell_L}\}$ again on the other side of R_j and following it up to the boundary of $\widetilde{U}_j$ (see Figure 8); on this new line, the bound (6-16) again holds; moreover, this new line is closer to the real axis than the line $\{\operatorname{Im} E = -e^{-\ell_L}\}$.

Let us call $\mathscr{C}_\ell$ the path obtained by gluing together the paths constructed in points (1)–(3) for $j^- \le j \le j^+$ and the half-lines $\frac{1}{2}(\lambda_{j^-} + \lambda_{j^- -1}) - i[e^{-\ell_L}, +\infty)$ and $\frac{1}{2}(\lambda_{j^+} + \lambda_{j^+ +1}) - i[e^{-\ell_L}, +\infty)$ (see (6-15)). One can then apply Rouché's theorem to compare the $+$-equations to the equations (2-4) and (2-8): by (6-14) and (6-16), on the line $\mathscr{C}_\ell$ one has $|S_L^-| < |S_L^+ + e^{-i\theta}|$ and

$$|\det(\Gamma_L(E) + e^{-i\theta(E)}) \det(\Gamma_L^+(E) + e^{-i\theta(E)})| \le \tfrac{1}{2}|\det(\Gamma_L(E) + e^{-i\theta(E)})|.$$

Thus, the number of solutions to equations (2-4) and (2-8) below the line $\mathscr{C}_\ell$ is bounded by $C\ell_L$; as $\mathscr{C}_\ell$ lies above $\{\operatorname{Im} E = -e^{-\ell_L}\}$, in the half-plane $\{\operatorname{Im} E < -e^{-\ell_L}\}$ the equations (2-4) and (2-8) have at most $C\ell_L$ solutions. We have proved Theorem 1.17.

6B6. *The proof of Theorem 1.18.* The first point in Theorem 1.18 is proved in the same way as point (2) in Theorem 1.14 up to the change of scales, L being replaced by ℓ_L. Pick scales $(\ell'_L)_L$ satisfying (1-22) such that $\ell'_L \ll \ell_L$. One has:

Lemma 6.7. *Fix two sequences $(a_L)_L$ and $(b_L)_L$ such that $a_L < b_L$. With probability one, for L sufficiently large,*

$$\begin{aligned}\#\{\text{eigenvalue of } H_{\omega,\ell_L - 2\ell'_L/\rho} \text{ in } [a_L + e^{-\ell'_L}, b_L - e^{-\ell'_L}]\} \\ \le \#\{\text{eigenvalue of } H_{\omega,L} \text{ in } [a_L, b_L] \text{ with localization center in } [\![0, \ell_L]\!]\} \\ \le \#\{\text{eigenvalue of } H_{\omega,\ell_L + 2\ell'_L/\rho} \text{ in } [a_L - e^{-\ell'_L}, b_L + e^{-\ell'_L}]\},\end{aligned}$$

where ρ is given by Lemma 6.2.

Proof. To prove Lemma 6.7, we apply Lemma 6.2 to $L = \ell_L + \ell'_L$ (i.e., for the operator H_ω restricted to the interval $[\![0, \ell_L + \ell'_L]\!]$) and $l_L = \ell'_L$. The probability of the bad set is the $O(L^{-\infty})$, thus summable in L. Using the localization estimate (6-3), one proves that

- each eigenvalue of $H_{\omega,\ell_L - 2\ell'_L/\rho}$ is at a distance of at most $e^{-\ell'_L}$ of an eigenvalue of $H_{\omega,L}$ with localization center in $[\![0, \ell_L]\!]$;
- each eigenvalue of $H_{\omega,L}$ with localization center in $[\![0, \ell_L]\!]$ is at a distance of at most $e^{-\ell'_L}$ of an eigenvalue of $H_{\omega,\ell_L + 2\ell'_L/\rho}$.

Lemma 6.7 follows. □

The first point in Theorem 1.18 is then Theorem 1.14(2) for the operators $H_{\omega,\ell_L - 2\ell'_L/\rho}$ and $H_{\omega,\ell_L + 2\ell'_L/\rho}$ and the fact that $\ell'_L \ll \ell_L$.

The proof of the second statement in Theorem 1.18 is very similar to that of Theorem 1.15. Fix a compact interval I in Σ°. As ℓ_L satisfies (1-22), one can find $\ell'_L < \ell''_L$ also satisfying (1-22) such that $e^{-\ell''_L} \ll e^{-\ell_L} \ll e^{-\ell'_L}$. For the same reasons as in the proof of Theorem 1.15, after rescaling all the resonances in $I - i(-\infty, 0)$ outside the strip $I - i[e^{-\ell'_L}, e^{-\ell''_L})$ are then pushed to either 0 or $i\infty$ as $L \to +\infty$.

On the other hand, the resonances in the strip $I - i[e^{-\ell'_L}, e^{-\ell''_L})$ are described by (3-19). The rescaled real and imaginary parts of the resonances (see (1-24)) now become $x_j = (E_{j,\omega} - E)\ell_L + o(1)$ and $y_j = \rho(E) d_{j,\omega}/\ell_L + o(1)$.

Now, to compute the limit of $\mathbb{P}(\#\{j \mid x_j \in I,\ y_j \in J\} = k)$, using the exponential decay property (6-3) it suffices to use [Germinet and Klopp 2014, Theorem 1.14]. Let us note here that [Germinet and Klopp 2014, Condition (1.50)] on the scales $(\ell_L)_L$ is slightly stronger than (1-22). That condition (1-22) suffices is a consequence of the stronger localization property known in the present case (compare Theorem 6.4 to [Germinet and Klopp 2014, Assumption (Loc)]). This completes the proof of the second point in Theorem 1.18. The final statement in 1.18 is proved in exactly the same way as Theorem 1.16.

The proof of Theorem 1.18 is complete.

6B7. *The proofs of Proposition 1.20 and Theorem 1.21.* Localization for the operator $H_\omega^{\mathbb{N}}$ can be described by the following:

Lemma 6.8. *There exists $\rho > 0$ and $q > 0$ such that, ω-almost surely, there exists $C_\omega > 0$ such that, for L sufficiently large, if*

(1) *$\varphi_{j,\omega}$ is a normalized eigenvector of $H_{\omega,L}$ associated to $E_{j,\omega}$ in Σ,*

(2) *$x_j(\omega) \in \mathbb{N}$ is a maximum of $x \mapsto |\varphi_{j,\omega}(x)|$ in $\mathbb{N}$,*

then, for $x \in \mathbb{N}$, one has

$$|\varphi_{j,\omega}(x)| \le C_\omega (1 + |x_j(\omega)|^2)^{q/2} e^{-\rho|x - x_j(\omega)|}. \tag{6-17}$$

Moreover, the mapping $\omega \mapsto C_\omega$ is measurable and $\mathbb{E}(C_\omega) < +\infty$.

This result for our model is a consequence of Theorem 6.1 (see, e.g., [Kunz and Souillard 1980; Cycon et al. 1987; Carmona and Lacroix 1990]) and [Germinet and Klopp 2014, Theorem 6.1].

We thus obtain the representation for the function Ξ_ω

$$\Xi_\omega(E) = \sum_j \frac{|\varphi_{j,\omega}(0)|^2}{E_{j,\omega} - E} + e^{-i \arccos(E/2)}. \tag{6-18}$$

As $H_\omega^{\mathbb{N}}$ satisfies a Dirichlet boundary condition at -1, one has

$$|\varphi_{j,\omega}(0)| > 0 \quad \text{for all } j \qquad \text{and} \qquad \sum_j |\varphi_{j,\omega}(0)|^2 = 1. \tag{6-19}$$

As $E \to -i\infty$, the representation (6-18) yields

$$\Xi_\omega(E) = -E^{-2} \sum_j |\varphi_{j,\omega}(0)|^2 E_{j,\omega} + O(E^{-3}) = -E^{-2} \langle \delta_0, H_\omega^{\mathbb{N}} \delta_0 \rangle + O(E^{-3}) = -\omega_0 E^{-2} + O(E^{-3}).$$

This proves the first point in Proposition 1.20.

As a direct consequence of Theorem 6.1 and the computation leading to Theorem 5.2 (see Section 5A2), we obtain that there exists $\tilde{c} > 0$ such that, for L sufficiently large, with probability at least $1 - e^{-\tilde{c}L}$ one has

$$\sup_{\operatorname{Im} E \le -e^{-\tilde{c}L}} \left| \int_{\mathbb{R}} \frac{dN_\omega(\lambda)}{\lambda - E} - \langle \delta_0, (H_{\omega,L} - E)^{-1} \delta_0 \rangle \right| \le e^{-\tilde{c}L}. \tag{6-20}$$

Taking

$$L = L_\varepsilon \sim c^{-1} |\log \varepsilon| \tag{6-21}$$

for some sufficiently small $c > 0$, this and Rouché's theorem implies that, with probability $1-\varepsilon^3$, the number of zeros of Ξ_ω (counted with multiplicity) in $I + i(-\infty, \varepsilon]$ is bounded

- from above by the number of resonances of H_{ω,L_ε} in $I_\varepsilon^+ + i(-\infty, -\varepsilon-\varepsilon^2]$,
- from below by the number of resonances of H_{ω,L_ε} in $I_\varepsilon^- + i(-\infty, -\varepsilon+\varepsilon^2]$,

where $I_\varepsilon^+ = [a-\varepsilon, b+\varepsilon]$ and $I_\varepsilon^+ = [a+\varepsilon, b-\varepsilon]$ if $I = [a, b]$.

Here, to apply Rouché's theorem, we apply the same strategy as in the proof of Theorem 1.17 and construct a path bounding a region larger (resp. smaller) than $I_\varepsilon^+ + i(-\infty, -\varepsilon-\varepsilon^2]$ (resp. $I_\varepsilon^- + i(-\infty, -\varepsilon+\varepsilon^2]$) on which one can guarantee $|S_L(E) + e^{-i\theta(E)}| \gtrsim 1$.

Now, we choose the constant c (see (6-21)) to be so small that $c < \min_{E\in I} \rho(E)$. Applying point (3) of Theorem 1.14 to H_{ω,L_ε} with this constant c, we obtain that the number of resonances of H_{ω,L_ε} in $I_\varepsilon^+ + i(-\infty, \varepsilon-\varepsilon^2]$ (resp. $I_\varepsilon^- + i(-\infty, \varepsilon+\varepsilon^2]$) is bounded from above (resp. bounded from below) by

$$\begin{aligned} L_\varepsilon \int_I \min\left(\frac{c}{\rho(E)}, 1\right) n(E)\, dE\, (1+O(1)) &= \frac{|\log \varepsilon|}{c} \int_I \frac{c}{\rho(E)} n(E)\, dE\, (1+O(1)) \\ &= |\log\varepsilon| \int_I \frac{n(E)}{\rho(E)}\, dE\, (1+O(1)). \end{aligned}$$

Hence, we obtain the second point of Proposition 1.20. The last point of this proposition is then an immediate consequence of the arguments developed to obtain the second point if one takes into account the following facts:

- The minimal distance between the Dirichlet eigenvalues of $H^{\mathbb{N}}_{\omega,L}$ is bounded from below by L^{-4} (see (6-10)).
- The growth of the function $E \mapsto S_L(E) + e^{-i\theta(E)}$ near the resonances (i.e., its zeros) is well controlled by Proposition 3.4.

Indeed, this implies that the resonances of $H^{\mathbb{N}}_{\omega,L}$ are simple in $I + i[-e^{-\sqrt{L}}, 0)$ (one can choose larger rectangles) and that near each resonance one can apply Rouché's theorem to control the zero of Ξ_ω. Note that this also yields, ω-almost surely, there exists c_ω such that

$$\min_{\substack{z \text{ zero of } \Xi_\omega \\ z\in I+i(-\varepsilon_\omega,0)}} \ \inf_{0<r<\varepsilon_\omega(\operatorname{Im} z)^{3/2}} \ \min_{|E-z|=r} \frac{|\Xi_\omega(E)|}{r} \gtrsim 1. \tag{6-22}$$

This completes the proof of Proposition 1.20.

Theorem 1.21 is a consequence of the following:

Theorem 6.9. *There exists $\tilde{c} > 0$ such that, ω-almost surely, for $L \ge 1$ sufficiently large one has*

$$\sup_{\substack{\operatorname{Re} E\in I \\ \operatorname{Im} E<-e^{-\tilde{c}L}}} \left| \Gamma_{L,\omega,\tilde\omega}(E) - \begin{pmatrix} \int_{\mathbb{R}} 1/(\lambda-E)\, dN_{\tilde\omega}(\lambda) & 0 \\ 0 & \int_{\mathbb{R}} 1/(\lambda-E)\, dN_\omega(\lambda) \end{pmatrix} \right| + \left| S_{L,\omega}(E) - \int_{\mathbb{R}} \frac{dN_\omega(\lambda)}{\lambda - E} \right| \le e^{-\tilde{c}L},$$

where $\Gamma_{L,\omega,\tilde\omega}(E)$ (resp. $S_{L,\omega}(E)$) is the matrix $\Gamma_L(E)$ (resp. the function $S_L(E)$)—see (2-9)*—constructed from the Dirichlet data on $[\![0, L]\!]$ of $-\Delta + V^{\mathbb{Z}}_{\omega,\tilde\omega,L}$ (resp. $-\Delta + V^{\mathbb{N}}_{\omega,L}$) (see* (1-26)*) using formula* (2-9) *(resp.* (2-4)*).*

Theorem 6.9 is proved exactly as Theorem 5.2 except that one uses the localization estimates (6-2) instead of the Combes–Thomas estimates.

Theorem 1.21 is then an immediate consequence of the estimate (6-20). Indeed, this implies that if z is a resonance for, e.g., $H^N_{\omega,L}$ in $I+i(-\infty, e^{\tilde{c}L}]$, then $|\Xi_\omega(z)| \le e^{-\tilde{c}L}$. By the last point of Proposition 1.20, ω-almost surely we know that the multiplicity of the zeros of Ξ_ω is bounded by N_ω. Moreover, for the zeros of Ξ_ω in $I+i(-\varepsilon_\omega, 0)$, we know the bound (6-22). This bound and (6-20) imply that

$$\max_{\substack{z \text{ zero of } \Xi_\omega \\ z\in I+i(-\varepsilon_\omega, e^{-\tilde{c}L})}} \max_{|E-z|=e^{-\tilde{c}L}} \frac{|\Xi_\omega(E)-(S_{\omega,L}(E)+e^{-i\theta(E)})|}{|\Xi_\omega(E)|} < e^{-\tilde{c}L}.$$

This yields Theorem 1.21(2) by an application of Rouché’s theorem. Point (1) is obtained in the same way, using Proposition 3.4, which gives

$$\max_{\substack{z \text{ resonance of } H^{\mathbb{N}}_{\omega,L} \\ z\in I+i(-\varepsilon_\omega, e^{-\tilde{c}L})}} \max_{|E-z|=e^{-\tilde{c}L}} \frac{|\Xi_\omega(E)-(S_{\omega,L}(E)+e^{-i\theta(E)})|}{|S_{\omega,L}(E)+e^{-i\theta(E)}|} < e^{-\tilde{c}L}.$$

The case of $H^{\mathbb{Z}}_{\omega,\tilde{\omega},L}$ is dealt with in the same way.

This completes the proof of Theorem 1.21.

6C. ***Estimates on the growth of eigenfunctions.*** In the present section we are going to prove Theorems 6.4 and 6.5. At the end of the section, we also prove the simpler Lemma 6.3.

The proof of Theorem 6.4 relies on locally uniform estimates on the rate of growth of the cocycle (1-15) attached to the Schrödinger operator, which we present now. Define

$$T_L(E,\omega) = T(E,\omega_L)\cdots T(E,\omega_0), \tag{6-23}$$

where

$$T(E,\omega_j) = \begin{pmatrix} E-\omega_j & -1 \\ 1 & 0 \end{pmatrix}.$$

We start with an upper bound on the large deviations of the growth rate of the cocycle that is uniform in energy. Fix $\alpha > 1$ and $\delta \in (0,1)$. For one part, the proof of Theorem 6.4 relies on the following:

Lemma 6.10. *Let $I \subset \mathbb{R}$ be a compact interval. For any $\delta > 0$, there exists $L_\delta > 0$ and $\eta > 0$ such that, for $L \ge L_\delta$ and any $K > 0$, one has*

$$\mathbb{P}\left(\frac{\log\|T_L(E;\tau^k(\omega))u\|}{L+1} \le \rho(E)+\delta \text{ for all } 0\le k\le K,\ E\in I,\ \|u\|=1\right) \ge 1-Ke^{-\eta(L+1)}, \tag{6-24}$$

where we recall that $\tau : \Omega \to \Omega$ denotes the left shift (i.e., if $\omega = (\omega_n)_{n\ge 0}$ then $[\tau(\omega)]_n = \omega_{n+1}$ for $n \ge 0$) and $\tau^n = \tau\circ\cdots\circ\tau$ n times.

At the heart of this result is a large deviation principle for the growth rate of the cocycle (see [Bougerol and Lacroix 1985, Section I and Theorem 6.1]). As it also serves in the proof of Theorem 6.4, we recall it now. One has:

Lemma 6.11. *Let $I \subset \mathbb{R}$ be a compact interval. For any $\delta > 0$, there exists $L_\delta > 0$ and $\eta > 0$ such that, for $L \geq L_\delta$, one has*

$$\sup_{\substack{E\in I\\ \|u\|=1}} \mathbb{P}\left(\left|\frac{\log\|T_L(E;\omega)u\|}{L+1} - \rho(E)\right| \geq \delta\right) \leq e^{-\eta(L+1)}. \tag{6-25}$$

While this result is not stated as is in [Bougerol and Lacroix 1985], it can be obtained from their Lemma 6.2 and Theorem 6.1. Indeed, by inspecting the proof of these results, it is clear that the quantities involved (in particular, the principal eigenvalue of $T(z;E) = T(z)$ in [loc. cit., Theorem 4.3]) are continuous functions of the energy E. Thus, taking this into account, the proof of [loc. cit., Theorem 6.1] yields, for our cocycle, a convergence that is locally uniform in energy, that is, (6-25).

To prove Theorem 6.4, in addition to Lemma 6.10 we also need to guarantee a uniform lower bound on the growth rate of the cocycle. We need this bound at least on the spectrum of $H_{\omega,L}$ with a good probability. Actually, this is the best one can hope for: a uniform bound in the style of (6-24) will not hold.

We prove:

Lemma 6.12. *Fix I a compact interval and $\delta > 0$. Pick $u \in \mathbb{C}^2$ with $\|u\| = 1$. For $0 \leq j \leq L$, if $j \leq L-1$, define*

$$\mathcal{K}_j^+(\omega, L, \delta, u) := \left\{E \in I \;\middle|\; \left|\frac{\log\|T_{L-(j+1)}^{-1}(E, \tau^{j+1}(\omega))u\|}{L-j} - \rho(E)\right| > \delta\right\}$$

and, if $1 \leq j$, define

$$\mathcal{K}_j^-(\omega, L, \delta, u) := \left\{E \in I \;\middle|\; \left|\frac{\log\|T_{j-1}(E,\omega)u\|}{j} - \rho(E)\right| > \delta\right\};$$

finally, define $\mathcal{K}_L^+(\omega, L, \delta, u) = \varnothing = \mathcal{K}_0^-(\omega, L, \delta, u)$.

Recall that $(E_{j,\omega})_{0\leq j\leq L}$ are the eigenvalues of $H_{\omega,L}$ and let $x_{j,\omega}$ be the associated localization centers.

For $0 \leq \ell \leq L$, define the sets

$$\Omega_B^+(L, \ell, \delta, u) := \{\omega \mid L - x_{j,\omega} \geq \ell \text{ and } E_{j,\omega} \in \mathcal{K}^+_{x_{j,\omega}}(\omega, L, \delta, u) \text{ for some } j\}$$

and

$$\Omega_B^-(L, \ell, \delta, u) := \{\omega \mid x_{j,\omega} \geq \ell \text{ and } E_{j,\omega} \in \mathcal{K}^-_{x_{j,\omega}}(\omega, L, \delta, u) \text{ for some } j\}.$$

Then the sets $\Omega_B^\pm(L, \ell, \delta, u)$ are measurable and, for any $\delta > 0$, there exist $\eta > 0$ and $\ell_0 > 0$ such that, for $L \geq \ell \geq \ell_0$, one has

$$\max\big(\mathbb{P}(\Omega_B^+(L,\ell,\delta,u)),\ \mathbb{P}(\Omega_B^-(L,\ell,\delta,u))\big) \leq \frac{(L+1)|I|e^{-\eta(\ell-1)}}{1-e^{-\eta}}. \tag{6-26}$$

Here, the constant η is the one given by (6-25).

First, let us explain the meaning of Lemma 6.12. Since by Lemma 6.10 we already control the growth of the cocycle from above, we see that in the definitions of the sets $\mathcal{K}_j^-(\omega, L, \delta, u)$ and $\mathcal{K}_j^+(\omega, L, \delta, u)$ it would have sufficed to require

$$\frac{\log \|T_{j-1}(E,\omega)u\|}{j} - \rho(E) \le -\delta \quad \text{and} \quad \frac{\log \|T^{-1}_{L-(j+1)}(E,\tau^{j+1}(\omega))u\|}{L-(j+1)} - \rho(E) \le -\delta,$$

respectively.

Hence, what Lemma 6.12 measures is that the probability that the cocycle at energy $E_{n,\omega}$ leading from a localization center $x_{n,\omega}$ to either 0 or L decays at a rate smaller than the rate predicted by the Lyapunov exponent.

The sets $\Omega^{\pm}_B(L,\ell,\delta,u)$ are the sets of bad configurations, i.e., the events when the rate of decay of the solution is far from the Lyapunov exponent. Indeed, for ω outside $\Omega^{\pm}_B(L,\ell,\delta)$, i.e., if the reverse of the inequalities defining $\mathcal{K}^{\pm}_j(\omega,L,\delta,u)$ hold, when $j = x_{n,\omega}$ and $E = E_{n,\omega}$ we know that the eigenfunction $\varphi_{n,\omega}$ has to decay from the center of localization $x_{n,\omega}$ (which is a local maximum of its modulus) towards the edges of the intervals at a rate larger than $\gamma(E_{n,\omega}) - \delta$. The eigenfunction being normalized, at the localization center it is of size at least $L^{-1/2}$. This will entail the estimates (6-4) and (6-5) with a good probability.

There is a major difference in the uniformity in energy obtained in Lemmas 6.12 and 6.10. In Lemma 6.12, we do not get a lower bound on the decay rate that is uniform all over I: it is merely uniform over the spectrum inside I (which is sufficient for our purpose, as we shall see). The reason for this difference in the uniformity between Lemma 6.10 and 6.12 is the same that makes the Lyapunov exponent $E \mapsto \rho(E)$ in general only upper semicontinuous and not lower semicontinuous (in the present situation, it actually is continuous).

We postpone the proofs of Lemmas 6.10 and 6.12 to the end of this section and turn to the proofs of Theorems 6.4 and 6.5.

6C1. *The proof of Theorem 6.4.* By Lemma 6.10, as $T_L(E,\omega) \in \mathrm{SL}(2,\mathbb{R})$, with probability at least $1 - KLe^{-\eta(L+1)}$, for $L \ge L_\delta$ and any $K > 0$, one also has

$$\forall 0 \le k \le K \ \ \forall E \in I \ \ \forall \|u\| = 1 \quad \frac{\log \|T_L^{-1}(E;\tau^k(\omega))u\|}{L+1} \le \rho(E) + \delta.$$

Now pick $\ell = C\log L$, where $C > 0$ is to be chosen later on. We know that, with probability $\mathbb{P}$ satisfying

$$\mathbb{P} \ge 1 - L^2 e^{-\eta\ell}, \tag{6-27}$$

for $L \ge L_\delta$, any $l \in [\ell, L]$ and any $k \in [0, L]$, one also has

$$\forall E \in I \ \ \forall \|u\| = 1 \ \ \frac{\log \|T_l^{-1}(E;\tau^k(\omega))u\|}{l+1} \le \rho(E) + \delta. \tag{6-28}$$

Let $\varphi_{j,\omega}$ be a normalized eigenfunction associated to the eigenvalue $E_{j,\omega} \in I$ with localization center $x_{j,\omega}$. By the definition of the localization center, one has

$$\frac{1}{L+1} \le \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\|^2 \le 1 \quad \text{and} \quad \frac{1}{L+1} \le \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}+1) \\ \varphi_{j,\omega}(x_{j,\omega}) \end{pmatrix} \right\|^2 \le 1. \tag{6-29}$$

By the eigenvalue equation, for $x \in [\![0, L]\!]$ one has

$$\begin{pmatrix} \varphi_{j,\omega}(x) \\ \varphi_{j,\omega}(x-1) \end{pmatrix} = \begin{cases} T_{x-x_{j,\omega}}(E;\tau^{x_{j,\omega}}(\omega)) \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} & \text{if } x \geq x_{j,\omega}, \\ T^{-1}_{x_{j,\omega}-x}(E;\tau^{x}(\omega)) \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} & \text{if } x \leq x_{j,\omega}. \end{cases} \tag{6-30}$$

Hence, by (6-24) and (6-28), with probability at least $1 - 2L^2 e^{-\eta\ell} - L^{-p}$, if $|x_{j,\omega} - x| \geq \ell$ then for $x_{j,\omega} < x \leq L$ one has

$$\begin{aligned} \frac{e^{-(\rho(E_{j,\omega})+\delta)|x-x_{j,\omega}|}}{\sqrt{L+1}} &\leq e^{-(\rho(E_{j,\omega})+\delta)|x-x_{j,\omega}|} \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\| \\ &\leq \left\| T_{x-x_{j,\omega}}(E;\tau^{x_{j,\omega}}(\omega)) \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\| = \left\| \begin{pmatrix} \varphi_{j,\omega}(x) \\ \varphi_{j,\omega}(x-1) \end{pmatrix} \right\| \end{aligned} \tag{6-31}$$

and for $0 \leq x < x_{j,\omega}$ one has

$$\begin{aligned} \left\| \begin{pmatrix} \varphi_{j,\omega}(x) \\ \varphi_{j,\omega}(x-1) \end{pmatrix} \right\| &= \left\| T^{-1}_{x-x_{j,\omega}}(E;\tau^{x_{j,\omega}}(\omega)) \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\| \\ &\geq e^{-(\rho(E_{j,\omega})+\delta)|x-x_{j,\omega}|} \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\| \geq \frac{e^{-(\rho(E_{j,\omega})+\delta)|x-x_{j,\omega}|}}{\sqrt{L+1}} \end{aligned} \tag{6-32}$$

On the other hand, by the definition of the Dirichlet boundary conditions, we know that

$$\begin{pmatrix} \varphi_{j,\omega}(0) \\ \varphi_{j,\omega}(-1) \end{pmatrix} = \varphi_{j,\omega}(0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \varphi_{j,\omega}(L+1) \\ \varphi_{j,\omega}(L) \end{pmatrix} = \varphi_{j,\omega}(L) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus,

$$\varphi_{j,\omega}(0) T_{x_{j,\omega}-1}(E;\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix}$$

and

$$\varphi_{j,\omega}(L) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = T_{L-x_{j,\omega}-1}(E;\tau^{x_{j,\omega}+1}(\omega)) \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}+1) \\ \varphi_{j,\omega}(x_{j,\omega}) \end{pmatrix}.$$

Thus, for $\omega \notin \Omega^+_B(L,\ell,\delta,u_+) \cup \Omega^-_B(L,\ell,\delta,u_-)$, where we have set $u_- := \binom{0}{1}$ and $u_+ := \binom{1}{0}$, we know that

$$e^{-(\rho(E_{j,\omega})-\delta)(L-x_{j,\omega})} \leq \left\| T^{-1}_{L-x_{j,\omega}-1}(E;\tau^{x_{j,\omega}+1}(\omega)) u_+ \right\| \quad \text{and} \quad e^{-(\rho(E_{j,\omega})-\delta)x_{j,\omega}} \leq \| T_{x_{j,\omega}-1}(E;\omega) u_- \|.$$

Thus we obtain that, for $\omega \notin \Omega^+_B(L,\ell,\delta,u_+) \cup \Omega^-_B(L,\ell,\delta,u_-)$, one has

$$|\varphi_{j,\omega}(L)| = \left\| T^{-1}_{L-x_{j,\omega}}(E;\tau^{x_{j,\omega}+1}(\omega)) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\|^{-1} \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}+1) \\ \varphi_{j,\omega}(x_{j,\omega}) \end{pmatrix} \right\| \leq e^{-(\rho(E_{j,\omega})-\delta)(L-x_{j,\omega}-1)} \tag{6-33}$$

and

$$|\varphi_{j,\omega}(0)| = \left\| T_{x_{j,\omega}}(E;\tau^{x_{j,\omega}}(\omega)) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\|^{-1} \left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\| \leq e^{-(\rho(E_{j,\omega})-\delta)(x_{j,\omega}-1)}. \tag{6-34}$$

The estimates given by Lemma 6.12 on the probability of $\Omega^+_B(L,\ell,\delta,u_+)$ and $\Omega^-_B(L,\ell,\delta,u_-)$ for $\ell = C\log L$ and the estimate (6-27) then imply that, with a probability at least $1-4L^2e^{-\eta(\ell-1)}-L^{-p}$, the bounds (6-31), (6-32), (6-33) and (6-34) hold. Thus, picking $\ell = C\log L$ for $C>0$ sufficiently large (depending only on η and, thus, on δ and p), these bounds hold with a probability at least $1-L^{-p}$. This completes the proof of Theorem 6.4.

Remark 6.13. One may wonder whether the uniform growth estimate given by Lemmas 6.10 and 6.12 is actually necessary in the proof of Theorem 6.4. That they are necessary is due to the fact that both the eigenvalue $E_{j,\omega}$ and the localization center $x_{j,\omega}$ (and, thus, the vector

$$\left\| \begin{pmatrix} \varphi_{j,\omega}(x_{j,\omega}) \\ \varphi_{j,\omega}(x_{j,\omega}-1) \end{pmatrix} \right\|$$

also) depend on ω. Thus, (6-25) is not sufficient to estimate the second term in the left-hand sides of (6-31) and (6-32).

6C2. *The proof of Theorem 6.5.* To prove Theorem 6.5, we follow the strategy that led to the proof of Theorem 6.4. First, note that (6-31) and (6-32) provide the expected lower bounds on the eigenfunction with the right probability. As for the upper bound, by (6-30), using the conclusions of Theorem 6.4 and the bounds given by Lemma 6.10, we know that, e.g., for $0\le x<x_{j,\omega}$,

$$\left\| \begin{pmatrix} \varphi_{j,\omega}(x) \\ \varphi_{j,\omega}(x-1) \end{pmatrix} \right\| = \left\| T_x(E;\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| |\varphi_{j,\omega}(0)| \le e^{(\rho(E_{j,\omega})+\delta)x} e^{-(\rho(E_{j,\omega})-\delta)x_{j,\omega}} \le e^{-(\rho(E_{j,\omega})-C\delta)|x-x_{j,\omega}|}$$

if $(1+C)x\le (C-1)x_{j,\omega}$, i.e., $2(1+C)^{-1}x_{j,\omega}\le x_{j,\omega}-x$.

For $x\ge x_{j,\omega}$ one reasons similarly and, thus, completes the proof of Theorem 6.5.

Remark 6.14. Actually, as the proof shows, the results one obtains are more precise than the claims made in Theorem 6.5 (see [Klopp $\ge$ 2016]).

6C3. *The proof of Lemma 6.12.* The proofs for the two sets $\Omega^\pm_B(L,\ell,\delta,u)$ are the same. We will only write out the one for $\Omega^+_B(L,\ell,\delta,u)$. Let us first address the measurability issue for $\Omega^+_B(L,\ell,\delta,u)$. The functions $\omega\mapsto E_{j,\omega}$ and $\omega\mapsto\varphi_{j,\omega}$ are continuous (as the eigenvalues and eigenvectors of finite-dimensional matrices depending continuously on the parameter $\omega=(\omega_j)_{0\le j\le L}$). Thus, for fixed j, the sets $\{\omega \mid E_{j,\omega}\in\mathcal{K}^-_j(\omega,L,\delta,u)\}$ and $\{\omega \mid x_{j,\omega}>j\}$ are open (we used the definition of $x_{j,\omega}$ as the leftmost localization center (see Theorem 6.4)). This yields the measurability of $\Omega^+_B(L,\ell,\delta,u)$.

We claim that

$$\frac{1}{L+1}\mathbf{1}_{\Omega^+_B(L,\ell,\delta,u)} \le \sum_{j=0}^{L+1-\ell} \langle \delta_j, \mathbf{1}_{\mathcal{K}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_j\rangle, \tag{6-35}$$

where $\mathbf{1}_{\mathcal{K}^+_j(\omega,L,\delta,u)}(H_{\omega,L})$ denotes the spectral projector associated to $H_{\omega,L}$ on the set $\mathcal{K}^+_j(\omega,L,\delta,u)$. Indeed, if one has $E_{j,\omega}\notin\mathcal{K}^+_{x_{j,\omega}}(\omega,L,\delta,u)$ for all j then the left-hand side of (6-35) vanishes and the right-hand side is nonnegative. On the other hand, if, for some j, one has $0\le x_{j,\omega}\le L-\ell$ and

$E_{j,\omega} \in \mathcal{H}^+_{x_{j,\omega}}(\omega, L, \delta, u)$, then we compute

$$\sum_{l=0}^{L-\ell} \langle \delta_l, \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_l\rangle = \sum_{l=0}^{L-\ell} \sum_{\substack{k \\ E_{k,\omega}\in\mathcal{H}^+_j(\omega,L,\delta,u)}} |\varphi_{k,\omega}(l)|^2 \geq |\varphi_{j,\omega}(x_{j,\omega})|^2 \geq \frac{1}{L+1} \geq \frac{1}{L+1}\mathbf{1}_{\Omega^+_B(L,\ell,\delta,u)}$$

by the definition of $x_{j,\omega}$.

An important fact is that, by construction (see Lemma 6.12), the set of energies $\mathcal{H}^+_j(\omega, L, \delta, u)$ does not depend on ω_j. Hence, denoting by $\mathbb{E}_{\omega_j}(\,\cdot\,)$ the expectation with respect to ω_j and $\mathbb{E}_{\hat{\omega}_j}(\,\cdot\,)$ the expectation with respect to $\hat{\omega}_j = (\omega_k)_{k\neq j}$, we compute

$$\mathbb{E}\biggl(\sum_{j=0}^{L-\ell}\langle \delta_j, \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_j\rangle\biggr) = \sum_{j=0}^{L-\ell} \mathbb{E}_{\hat{\omega}_j}\bigl(\mathbb{E}_{\omega_j}\bigl(\langle \delta_j, \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_j\rangle\bigr)\bigr).$$

As ω_j is assumed to have a bounded, compactly supported distribution and as $\mathcal{H}^+_j(\omega, L, \delta, u)$ does not depend on ω_j, a standard spectral averaging lemma (see, e.g., [Simon 2005, Theorem 11.8]) yields

$$\mathbb{E}_{\omega_j}\bigl(\langle \delta_j, \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_j\rangle\bigr) \leq |\mathcal{H}^+_j(\omega, L, \delta, u)|,$$

where $|\cdot|$ denotes the Lebesgue measure. Thus, we obtain

$$\mathbb{E}\biggl(\sum_{j=0}^{L-\ell}\langle \delta_j, \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(H_{\omega,L})\delta_j\rangle\biggr) \leq \sum_{j=0}^{L-\ell} \mathbb{E}_{\hat{\omega}_j}\bigl(|\mathcal{H}^+_j(\omega, L, \delta, u)|\bigr) = \sum_{j=0}^{L-\ell} \mathbb{E}\bigl(|\mathcal{H}^+_j(\omega, L, \delta, u)|\bigr). \tag{6-36}$$

By Lemma 6.11 and the Fubini–Tonelli theorem, we know that

$$\begin{aligned}\mathbb{E}\bigl(|\mathcal{H}^+_j(\omega, L, \delta, u)|\bigr) &= \mathbb{E}\biggl(\int_I \mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(E)\,dE\biggr) = \int_I \mathbb{E}(\mathbf{1}_{\mathcal{H}^+_j(\omega,L,\delta,u)}(E))\,dE \\ &\leq |I|\sup_{E\in I}\mathbb{P}\biggl(\biggl|\frac{\log\|T^{-1}_{L-(j+1)}(E,\omega)u\|}{L-j} - \rho(E)\biggr| > \delta\biggr) \\ &\leq |I|r^{-\eta(L-j)}.\end{aligned}$$

Taking the expectation of both sides of (6-35) and plugging this into (6-36), we obtain

$$\mathbb{P}(\Omega^+_B(L, \ell, \delta, u)) \leq (L+1)|I|e^{-\eta(\ell-1)}\sum_{j=0}^{L-\ell} e^{-\eta j} \leq \frac{(L+1)|I|e^{-\eta(\ell-1)}}{1-e^{-\eta}}.$$

In the same way, one obtains

$$\mathbb{P}(\Omega^-_B(L, \ell, \delta, u)) \leq \frac{(L+1)|I|e^{-\eta(\ell-1)}}{1-e^{-\eta}}.$$

This completes the proof of Lemma 6.12.

Remark 6.15. This proof can be seen as the analogue for products of finitely many random matrices of the so-called Kotani trick (see, e.g., [Cycon et al. 1987]).

6C4. *The proof of Lemma 6.10.* The basic idea of this proof is to use the estimate (6-25), in particular, the exponentially small probability and some perturbation theory for the cocycles so as to obtain a uniform estimate.

Let η be given by (6-25). Fix $\eta' < \frac{1}{2}\eta$ and write

$$I = \bigcup_{j\in J}[E_j, E_{j+1}], \quad \text{where} \quad \tfrac{1}{2}e^{-\eta'(L+1)} \le E_{j+1} - E_j \le 2e^{-\eta'(L+1)}; \tag{6-37}$$

thus, $\#J \lesssim e^{\eta'(L+1)}$.

We now want to estimate what happens for $E \in [E_j, E_{j+1}]$. Using (1-15) and

$$\begin{pmatrix} E - V_\omega(n) & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} E_j - V_\omega(n) & -1 \\ 1 & 0 \end{pmatrix} = (E - E_j)\Delta T, \quad \text{where} \quad \Delta T := \left| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|,$$

we compute

$$T_L(E, \omega) = T_L(E_j, \omega) + \sum_{l=1}^{L} (E - E_j)^l S_l, \tag{6-38}$$

where

$$\begin{aligned} S_l &:= \sum_{n_1 < n_2 < \cdots < n_l} T_{n_1}(E_j, \tau^{L-n_1}\omega) \times \Delta T \times T_{n_2-n_1-1}(E_j, \tau^{n_2}\omega) \times \Delta T \times \cdots \times \Delta T \times T_{L-n_l-1}(E_j, \tau^{n_l}\omega) \\ &= \sum_{n_1 < n_2 < \cdots < n_l} \prod_{m=2}^{l} \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, T_{n_m - n_{m-1}-1}(E_j, \tau^{n_m}\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \left| T_{n_1}(E_j, \tau^{L-n_1}\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| T_{L-n_l-1}(E_j, \tau^{n_l}\omega). \end{aligned}$$

Clearly, as the random variables have compact support, one has the uniform bound

$$\sup_{\substack{E\in I \\ \omega\in\Omega}} \|T_L(E;\omega)\| \le e^{CL}. \tag{6-39}$$

Thus one has

$$\sup_{\omega\in\Omega} \|S_l\| \le L^l e^{CL}. \tag{6-40}$$

Hence, for l_0 fixed, one computes

$$\left\| \sum_{l=l_0}^{L} (E - E_j)^l S_l \right\| \le \sum_{l=l_0}^{L} (E - E_j)^l \|S_l\| \le \sum_{l=l_0}^{L} e^{-\eta'(L+1)l} L^l e^{CL} \le 1 \tag{6-41}$$

if $\eta' l_0 > 2C$ and L is sufficiently large (depending only on η' and C).

We now assume that l_0 satisfies $\eta' l_0 > 2C$ and pick $1 \le l \le l_0$. Pick $\delta_0 \in (0, 1)$ small, to be fixed later. Assume moreover that L is such that $\delta_0 L \ge L_\delta$, where L_δ is as defined in Lemma 6.11. Then, by Lemma 6.11, for $m \in \{2, \dots, l\}$ one has

(1) either $n_m - n_{m-1} \le L_\delta$, in which case one has

$$\|T_{n_m - n_{m-1}-1}(E_j, \tau^{n_m - 1}\omega)\| \le e^{C(n_m - n_{m-1})};$$

(2) or $n_m - n_{m-1} \geq L_\delta$, in which case, by (6-25), with probability at least equal to $1 - e^{-\eta(n_m - n_{m-1})/2}$, one has

$$\|T_{n_m - n_{m-1} - 1}(E_j, \tau^{n_m - 1}\omega)\| \leq e^{(n_m - n_{m-1})(\rho(E_j)+\delta)}.$$

Define

$$G_{n_1,\dots,n_l} = \{m \in \{2, \dots, l\} \mid n_m - n_{m-1} \geq L_\delta\} \quad \text{and} \quad B_{n_1,\dots,n_l} = \{2, \dots, l\} \setminus G_{n_1,\dots,n_l}.$$

By definition, one has

$$\sum_{m \in B_{n_1,\dots,n_l}} (n_m - n_{m-1}) \leq l L_\delta \quad \text{and} \quad \sum_{m \in G_{n_1,\dots,n_l}} (n_m - n_{m-1}) \geq L - l L_\delta. \tag{6-42}$$

For a fixed sequence $n_1 < n_2 < \cdots < n_m$, the random variables $(T_{n_{m'} - n_{m'-1} - 1}(E_j, \tau^{n_{m'}}\omega))_{1 \leq m' \leq m}$ are independent. Hence, by (6-25), for a fixed $(m_1, \dots, m_K) \in G_{n_1,\dots,n_l}$, one has

$$\mathbb{P}\Big(\inf_{1 \leq k \leq K} \big\|T_{n_{m_k} - n_{m_{k-1}} - 1}(E_j, \tau^{n_{m_k}}\omega)\big\| \geq e^{(\rho(E_j)+\delta)(n_{m_k} - n_{m_{k-1}})}\Big) \leq e^{-\eta \sum_{k=1}^{K} n_{m_k} - n_{m_{k-1}}}.$$

Thus, for $\varepsilon \in (0, 1)$, one has

$$\mathbb{P}\Bigg(\inf_{1 \leq k \leq K} \|T_{n_{m_k} - n_{m_{k-1}} - 1}(E_j, \tau^{n_{m_k} - 1}\omega)\| \geq e^{(\rho(E_j)+\delta)(n_{m_k} - n_{m_{k-1}})} \\ \text{for some } (m_1, \dots, m_K) \in G_{n_1,\dots,n_l} \text{ with } \sum_{k=1}^{K} n_{m_k} - n_{m_{k-1}} \geq \varepsilon L\Bigg) \leq L^l e^{-\eta\varepsilon L}.$$

Hence, with probability at least $1 - L^l e^{-\eta\varepsilon L}$, we know that there exists $(m_1, \dots, m_K) \in G_{n_1,\dots,n_l}$ such that

$$\sum_{k=1}^{K} n_{m_k} - n_{m_{k-1}} \geq L - l L_\delta - \varepsilon L \quad \text{and} \quad \|T_{n_{m_k} - n_{m_{k-1}} - 1}(E_j, \tau^{n_{m_k} - 1}\omega)\| \leq e^{(\rho(E_j)+\delta)(n_{m_k} - n_{m_{k-1}})}$$

for all $1 \leq k \leq K$. Using the estimates (6-42) and (6-39) for the remaining terms in the product below, for any given m-tuple $(n_1, \dots, n_m)$ one obtains

$$\mathbb{P}\Bigg(\prod_{m=1}^{l} \|T_{n_m - n_{m-1} - 1}(E_j, \tau^{n_{m_k} - 1}\omega)\| \leq e^{(\rho(E_j)+\delta)(1-\varepsilon)(L - lL_\delta) + C(\varepsilon L + lL_\delta)}\Bigg) \geq 1 - L^l e^{-\eta\varepsilon L}.$$

Hence, with probability at least $1 - l_0 L^{l_0} e^{-\eta\varepsilon L}$, for $1 \leq l \leq l_0$ we estimate

$$\begin{aligned} \|S_l\| &\leq \sum_{n_1 < n_2 < \cdots < n_l} \prod_{m=1}^{l} \|T_{n_m - n_{m-1} - 1}(E_j, \tau^{n_{m_k}}\omega)\| \\ &\leq L^l e^{(\rho(E_j)+\delta)(1-\varepsilon)L + C\varepsilon L + (C - (\rho(E_j)+\delta)(1-\varepsilon)) l L_\delta} \\ &\leq L^l e^{[\rho(E_j)+\delta+(C-\rho(E_j)-\delta)\varepsilon]L + [C-(\rho(E_j)+\delta)(1-\varepsilon)]L_\delta l} \\ &\leq L^{l_0} e^{[\rho(E_j)+\delta+(C-\rho(E_j)-\delta)\varepsilon]L + [C-(\rho(E_j)+\delta)(1-\varepsilon)]L_\delta l_0}. \end{aligned} \tag{6-43}$$

It remains now to choose the quantities η', l_0 and ε so that the following requirements are satisfied:

$$\eta' l_0 > 2C, \quad (C-\rho(E_j)-\delta)\varepsilon \le \frac{\delta}{2}, \quad l_0 L^{l_0} e^{-\eta\varepsilon L} e^{\eta'(L+1)} \ll 1 \quad \text{and} \quad \frac{[C-(\rho(E_j)+\delta)(1-\varepsilon)]L_\delta l_0}{L+1} \le \frac{\delta}{2(\rho(E_j)+\delta)}. \tag{6-44}$$

Fixing ε small, picking $0<\eta'<\frac{1}{3}\eta\varepsilon$ and setting $l_0 = L^\alpha$, where $\alpha \in (0,1)$, we see that all the conditions in (6-44) are satisfied for L sufficiently large. Moreover, one has

$$l_0 L^{l_0} e^{-\eta\varepsilon L} e^{\eta'(L+1)} \le e^{-\eta\varepsilon L/2}.$$

Plugging this and the last estimate in (6-43) into (6-38), we obtain that, with probability at least $1-e^{-\eta\varepsilon L/2}$, for any $j\in J$ (see (6-37)) and $E\in[E_j, E_{j+1}]$ one has

$$\|T_L(E,\omega)-T_L(E_j,\omega)\| \le 1+\sum_{l=1}^{l_0} e^{-\eta' l(L+1)} L^l e^{(\rho(E_j)+2\delta)L} \le 1+e^{(\rho(E_j)+2\delta)(L+1)}. \tag{6-45}$$

As ρ is continuous (see, e.g., [Bougerol and Lacroix 1985]), one gets that, for any $\delta>0$ and L sufficiently large, with probability at least $1-e^{-\eta\varepsilon L/2}$, one has, for any $E\in I$,

$$\|T_L(E,\omega)\| \lesssim e^{(\rho(E)+2\delta)(L+1)}.$$

Hence, as $T_L(E,\omega)\in \mathrm{SL}(2,\mathbb{R})$, one has $\|T_L^{-1}(E,\omega)\| \lesssim e^{(\rho(E)+2\delta)(L+1)}$.

Using the fact that the probability measure on Ω is invariant under the shift (it is a product measure), we obtain (6-24). This completes the proof of Lemma 6.10.

6C5. *The proof of Lemma 6.3.* Assume the realization ω is such that the conclusions of Lemma 6.2 hold in I for the scales $l_L = 2\log L$. Fix $\alpha>0$ and let $\mathscr{E}_{L,\omega}$ be the set of indices of the eigenvalues $(E_{j,\omega})_{0\le j\le L}$ of $H_{\omega,L}$ having a localization center in $[\![L-\ell_L, L]\!]$. Fix $C>\alpha>0$ and consider the projector $\Pi_C := \mathbf{1}_{[\![L-C\ell_L,L]\!]}$ in $\ell^2([\![0,L]\!])$.

Consider the Gram matrices

$$G(\mathscr{E}_{L,\omega}) = \big((\langle\varphi_{j,\omega},\varphi_{j,\omega}\rangle)\big)_{(n,m)\in\mathscr{E}_{L,\omega}\times\mathscr{E}_{L,\omega}} = \mathrm{Id}_N,$$

where $N=\#\mathscr{E}_{L,\omega}$, and

$$G_\pi(\mathscr{E}_{L,\omega}) = \big((\langle\Pi_C\varphi_{j,\omega},\Pi_C\varphi_{j,\omega}\rangle)\big)_{(n,m)\in\mathscr{E}_{L,\omega}\times\mathscr{E}_{L,\omega}}.$$

By definition, the rank of $G_\pi(\mathscr{E}_{L,\omega})$ is bounded by the rank of Π_C, i.e., by $C\ell_L$. Moreover, as by (6-3) one has $\|(1-\Pi_C)\varphi_{j,\omega}\| \le L^q e^{-\rho\eta C\ell_L}$, one has

$$\|\mathrm{Id}_N - G_\pi(\mathscr{E}_{L,\omega})\| \le L^{2+q}e^{-\rho\eta C\ell_L} \le L^{2+q-C\rho\eta}.$$

Thus, picking $C\eta\rho > q+2$ yields that, for L sufficiently large, $G_\pi(\mathscr{E}_{L,\omega})$ is invertible and its rank is N. This yields $\#\mathscr{E}_{L,\omega} = N \le C\ell_L$ and the proof of Lemma 6.3 is complete.

6D. ***The half-line random perturbation*****:** ***the proof of Theorem 1.23.*** Using the same notations as in Section 5C, we can write

$$H^\infty = \begin{pmatrix} H^-_{\omega,-1} & |\delta_{-1}\rangle\langle\delta_0| \\ |\delta_0\rangle\langle\delta_{-1}| & -\Delta_0^+ \end{pmatrix},$$

where

- $-\Delta_0^+$ is the Dirichlet Laplacian on $\ell^2(\mathbb{N})$,
- $H^-_{\omega,-1} = -\Delta + V_\omega$ on $\ell^2(\{n \le -1\})$ with Dirichlet boundary conditions at 0.

Define the operators

$$\Gamma_\omega(E) := -\Delta_0^+ - E - \langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle|\delta_0\rangle\langle\delta_0|,$$
$$\widetilde{\Gamma}_\omega(E) := H^-_{\omega,-1} - E - \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle|\delta_{-1}\rangle\langle\delta_{-1}|.$$

For $\operatorname{Im} E \neq 0$, the numbers $\langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle$ and $\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle$ have nonvanishing imaginary parts of the same sign; hence, the complex number

$$(\langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle)^{-1} - \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle$$

does not vanish. Thus, by rank-one perturbation theory (see, e.g., [Simon 2005]), we thus know that $\Gamma_\omega(E)$ and $\widetilde{\Gamma}_\omega(E)$ are invertible for $\operatorname{Im} E \neq 0$ and that

$$\Gamma_\omega^{-1}(E) = (-\Delta_0^+ - E)^{-1} + \frac{|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle\langle\delta_0|(-\Delta_0^+ - E)^{-1}|}{\left(\langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\right)^{-1} - \langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle} \tag{6-46}$$

$$\widetilde{\Gamma}_\omega^{-1}(E) = (H^-_{\omega,-1} - E)^{-1} + \frac{|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|}{\left(\langle\delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle\right)^{-1} - \langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle}. \tag{6-47}$$

Thus, for $\operatorname{Im} E \neq 0$, using Schur's complement formula we compute

$$(H^\infty_\omega - E)^{-1} = \begin{pmatrix} \widetilde{\Gamma}_\omega^{-1}(E) & \gamma(E) \\ \gamma^*(\bar{E}) & \Gamma_\omega^{-1}(E) \end{pmatrix}, \tag{6-48}$$

where $\gamma^*(\bar{E})$ is the adjoint of $\gamma(\bar{E})$ and

$$\gamma(E) := -\left|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\langle\delta_0|\Gamma_\omega^{-1}(E)\right|$$

6D1. *The continuation through* $(-2,2)\setminus\Sigma$. Let us start with the analytic continuation through $(-2,2)\setminus\Sigma$.

One easily checks that the function $E \mapsto \langle\delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle^{-1}$ is analytic outside Σ, the essential spectrum of $H^-_{\omega,-1}$, and has simple zeros at the isolated eigenvalues of $H^-_{\omega,-1}$. Hence, $E \mapsto \Gamma_\omega^{-1}(E)$ can be analytically continued near an isolated eigenvalue of $H^-_{\omega,-1}$ different from -2 and 2.

As for $\widetilde{\Gamma}_\omega^{-1}$, using the spectral decomposition of $(H^-_{\omega,-1} - E)^{-1}$, as for any eigenvector of $H^-_{\omega,-1}$, say φ, one has $\langle\delta_{-1},\varphi\rangle \neq 0$; for E_0 an isolated eigenvalue of $H^-_{\omega,-1}$ different from -2 and 2, doing a polar decomposition of $\widetilde{\Gamma}_\omega^{-1}$ near E_0 one checks that $E \mapsto \widetilde{\Gamma}_\omega^{-1}(E)$ can be analytically continued to a neighborhood of E_0.

Finally let us check what happens with γ. We compute

$$\gamma(E) = -\langle\delta_{-1}|(H^{-}_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle^{-1}\big|(H^{-}_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle\langle\delta_0|(-\Delta_0^{+}-E)^{-1}\big|.$$

As $E\mapsto\langle\delta_{-1}|(H^{-}_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle^{-1}(H^{-}_{\omega,-1}-E)^{-1}$ is analytic near any isolated eigenvalue of $H^{-}_{\omega,-1}$, we see that $E\mapsto\gamma(E)$ can be can be analytically continued to a neighborhood of an isolated eigenvalue of $H^{-}_{\omega,-1}$.

Hence, the representation (6-48) immediately shows that the resolvent $(H^{\infty}_{\omega}-E)^{-1}$ can be continued through $(-2,2)\setminus\Sigma$, the poles of the continuation being given by the zeros of the function

$$E\mapsto 1-\langle\delta_0|(-\Delta_0^{+}-E)^{-1}|\delta_0\rangle\langle\delta_{-1}|(H^{-}_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle = 1-e^{i\theta(E)}\int_{\mathbb{R}}\frac{dN_\omega(\lambda)}{\lambda-E}.$$

6D2. *No continuation through* $(-2,2)\cap\Sigma^{\circ}$. Let us study the analytic continuation through $(-2,2)\cap\Sigma^{\circ}$. Considering the lower right coefficient of this matrix, we see that, when coming from upper half-plane through $(-2,2)\cap\Sigma^{\circ}$, $E\mapsto(H^{\infty}_{\omega}-E)^{-1}$ can be continued meromorphically to the lower half plane (as an operator from $\ell^2_{\text{comp}}(\mathbb{Z})$ to $\ell^2_{\text{loc}}(\mathbb{Z})$) only if $E\mapsto\Gamma^{-1}_{\omega}(E)$ can be continued meromorphically (as an operator from $\ell^2_{\text{comp}}(\mathbb{N})$ to $\ell^2_{\text{loc}}(\mathbb{N})$).

As $E\mapsto(-\Delta_0^{+}-E)^{-1}$ can be analytically continued (see Section 2), by (6-46) the meromorphic continuation of $E\mapsto\Gamma^{-1}_{\omega}(E)$ will exist if and only if the complex-valued map

$$E\mapsto g_\omega(E) := \frac{1}{\big(\langle\delta_{-1}|(H^{-}_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle\big)^{-1}-\langle\delta_0|(-\Delta_0^{+}-E)^{-1}|\delta_0\rangle}$$

can be meromorphically continued from the upper half-plane through $(-2,2)\cap\Sigma^{\circ}$. Fix ω such that the spectrum of $H^{-}_{\omega,-1}$ is equal to Σ and pure point (this is almost sure; see, e.g., [Carmona and Lacroix 1990; Pastur and Figotin 1992]). As δ_{-1} is a cyclic vector for $H^{-}_{\omega,-1}$, for E an eigenvalue of $H^{-}_{\omega,-1}$ one then has

$$\lim_{\varepsilon\to0^{+}}\big(\langle\delta_{-1}|(H^{-}_{\omega,-1}-E-i\varepsilon)^{-1}|\delta_{-1}\rangle\big)^{-1}=0. \tag{6-49}$$

Hence, if the analytic continuation of g_ω would exist on $(-2,2)\cap\Sigma^{\circ}$ it would be equal to

$$g_\omega(E+i0) = -\frac{1}{\langle\delta_0|(-\Delta_0^{+}-E-i0)^{-1}|\delta_0\rangle}. \tag{6-50}$$

By analyticity of both sides, this in turn would imply that (6-50) holds on the whole upper half-plane; thus, in view of the definition of g_ω, that (6-49) holds on the whole upper half plane: this is absurd! Thus, we have proved that, ω-almost surely, $E\mapsto(H^{\infty}_{\omega}-E)^{-1}$ does not admit a meromorphic continuation through $(-2,2)\cap\Sigma^{\circ}$.

6D3. *Absolutely continuity of the spectrum of* H^{∞}_{ω} *in* $(-2,2)\cap\Sigma^{\circ}$. Let us now prove that the spectral measure of H^{∞}_{ω} in $(-2,2)\cap\Sigma^{\circ}$ is purely absolutely continuous. It suffices (see, e.g., [Teschl 2000, Section 2.5; Simon 2005, Theorem 11.6]) to prove that, for all $E\in(-2,2)\cap\Sigma^{\circ}$, one has

$$\limsup_{\varepsilon\to0^{+}}\big|\langle\delta_0,(H^{\infty}_{\omega}-E-i\varepsilon)^{-1}\delta_0\rangle\big|+\big|\langle\delta_{-1},(H^{\infty}_{\omega}-E-i\varepsilon)^{-1}\delta_{-1}\rangle\big|<+\infty.$$

Using (6-46), (6-47) and (6-48), for $\operatorname{Im} E \neq 0$ we compute

$$\langle \delta_{-1}, (H_\omega^\infty - E)^{-1}\delta_{-1}\rangle = \frac{\langle \delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle}{1 - \langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle \cdot \langle \delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle}, \tag{6-51}$$

$$\langle \delta_{-n}, (H_\omega^\infty - E)^{-1}\delta_{m}\rangle = \frac{-\langle \delta_{-n}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_m\rangle}{1 - \langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle \cdot \langle \delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle} \tag{6-52}$$

for $n \geq 1$ and $m \leq 0$, and

$$\langle \delta_0, (H_\omega^\infty - E)^{-1}\delta_0\rangle = \frac{\langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle}{1 - \langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle \cdot \langle \delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle}. \tag{6-53}$$

Thus, to prove the absolute continuity of the spectral measure of H_ω^∞ in $(-2,2)\cap\Sigma^\circ$, it suffices to prove that, for $E \in (-2,2)\cap\Sigma^\circ$, one has

$$\limsup_{\varepsilon\to 0^+}\left(\left|\frac{1}{\big(\langle \delta_{-1}|(H^-_{\omega,-1} - E - i\varepsilon)^{-1}|\delta_{-1}\rangle\big)^{-1} - \langle \delta_0|(-\Delta_0^+ - E - i\varepsilon)^{-1}|\delta_0\rangle}\right| + \left|\frac{1}{\big(\langle \delta_0|(-\Delta_0^+ - E - i\varepsilon)^{-1}|\delta_0\rangle\big)^{-1} - \langle \delta_{-1}|(H^-_{\omega,-1} - E - i\varepsilon)^{-1}|\delta_{-1}\rangle}\right|\right) < \infty.$$

This is the case, as

- the signs of the imaginary parts of $-\big(\langle \delta_{-1}|(H^-_{\omega,-1} - E - i\varepsilon)^{-1}|\delta_{-1}\rangle\big)^{-1}$ and $\langle \delta_0|(-\Delta_0^+ - E - i\varepsilon)^{-1}|\delta_0\rangle$ are the same (negative if $\operatorname{Im} E < 0$ and positive if $\operatorname{Im} E > 0$),
- for $E \in (-2,2)$, $\langle \delta_0|(-\Delta_0^+ - E - i\varepsilon)^{-1}|\delta_0\rangle$ has a finite limit when $\varepsilon \to 0^+$,
- for $E \in (-2,2)$, the imaginary part of $\langle \delta_0|(-\Delta_0^+ - E - i\varepsilon)^{-1}|\delta_0\rangle$ does not vanish in the limit $\varepsilon \to 0^+$.

So, we have proved the part of Theorem 1.23 concerning the absence of analytic continuation of the resolvent of H_ω^∞ through $(-2,2)\cap\Sigma^\circ$ and the nature of its spectrum in this set.

6D4. *The spectrum of H_ω^∞ is pure point in $\Sigma^\circ \setminus [-2,2]$.* Let us now prove the last part of Theorem 1.23. The proof relies again on (6-48). We pick $\beta \in \big(0, \frac{1}{2}\alpha\big)$, where α is determined by Theorem 6.1 for $H^-_{\omega,-1}$. Then, for $n \geq 1$ and $m \leq 0$, using the Cauchy–Schwartz inequality, for $\operatorname{Im} E \neq 0$ we compute

$$\begin{aligned}\mathbb{E}\Big(\big|\langle \delta_{-n}, (H_\omega^\infty - E)^{-1}\delta_m\rangle\big|^\beta\Big)^2 &\leq \big|\langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_m\rangle\big|^2 \cdot \mathbb{E}\Big(\big|\langle \delta_{-n}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\big|^{2\beta}\Big)\\ &\quad\cdot \mathbb{E}\left(\left|\frac{1}{1 - \langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_0\rangle \cdot \langle \delta_{-1}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle}\right|^{2\beta}\right).\end{aligned} \tag{6-54}$$

For a compact interval $J \subset (-2,2)\setminus\Sigma$, we know that, for $n \geq 1$ and $m \leq 0$,

- $\sup_{\operatorname{Im} E\neq 0}\big|\langle \delta_0|(-\Delta_0^+ - E)^{-1}|\delta_m\rangle\big| \lesssim e^{-cm}$ by the Combes–Thomas estimates;
- $\sup_{\operatorname{Im} E\neq 0}\mathbb{E}\Big(\big|\langle \delta_{-n}|(H^-_{\omega,-1} - E)^{-1}|\delta_{-1}\rangle\big|^{2\beta}\Big) \lesssim e^{-2\beta\rho n}$ by the characterization (6-1) of localization in Σ for $H^-_{\omega,-1}$.

It suffices now to estimate the last term in (6-54) using a standard decomposition of rank-one perturbations (see, e.g., [Simon 2005; Aizenman and Molchanov 1993]); one writes

$$\frac{1}{1-\langle\delta_0|(-\Delta_0^+-E)^{-1}|\delta_0\rangle\cdot\langle\delta_{-1}|(H^-_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle}=\frac{\omega_{-1}-b}{\omega_{-1}-a},$$

where a and b only depend on $(\omega_{-n})_{n\geq 2}$. Thus, as $(\omega_{-n})_{n\geq 1}$ have a bounded density, for $\operatorname{Im} E\neq 0$ one has

$$\mathbb{E}\left(\left|\frac{1}{1-\langle\delta_0|(-\Delta_0^+-E)^{-1}|\delta_0\rangle\cdot\langle\delta_{-1}|(H^-_{\omega,-1}-E)^{-1}|\delta_{-1}\rangle}\right|^{2\beta}\right)\leq\mathbb{E}_{(\omega_{-n})_{n\geq 2}}\mathbb{E}_{\omega_{-1}}\left(\left|\frac{\omega_{-1}-b}{\omega_{-1}-a}\right|^{2\beta}\right)\leq C_\beta<+\infty.$$

Thus, we have proved that, for a compact interval $J\subset\Sigma\setminus[-2,2]$, for $\beta\in\left(0,\frac{1}{2}\alpha\right)$ and some $\tilde\rho>0$, for $n\geq 1$ and $m\leq 0$ one has

$$\sup_{\substack{\operatorname{Im} E\neq 0\\ \operatorname{Re} E\in I}}\mathbb{E}\left(\left|\langle\delta_{-n},(H_\omega^\infty-E)^{-1}\delta_m\rangle\right|^\beta\right)<C_\beta e^{-\tilde\rho(m-n)}.$$

In the same way, using (6-51) and (6-53), one proves that

$$\sup_{\substack{\operatorname{Im} E\neq 0\\ \operatorname{Re} E\in I}}\mathbb{E}\left(\left|\langle\delta_0,(H_\omega^\infty-E)^{-1}\delta_0\rangle\right|^\beta+\left|\langle\delta_{-1},(H_\omega^\infty-E)^{-1}\delta_{-1}\rangle\right|^\beta\right)<+\infty.$$

Thus, we have proved that, for some $\tilde\rho>0$, one has

$$\sup_{\substack{\operatorname{Im} E\neq 0\\ \operatorname{Re} E\in I}}\sup_{m\in\mathbb{Z}}\mathbb{E}\left(\sum_{n\in\mathbb{Z}}e^{\tilde\rho(m-n)}\left|\langle\delta_{-n},(H_\omega^\infty-E)^{-1}\delta_m\rangle\right|^\beta\right)<+\infty.$$

Hence, we know that the spectrum of H_ω^∞ in $\Sigma\setminus[-2,2]$ (as J can be taken arbitrarily, contained in this set) is pure point associated to exponentially decaying eigenfunctions (see, e.g., [Aizenman and Molchanov 1993; Aizenman 1994; Aizenman et al. 2001]). This completes the proof of Theorem 1.23.

Appendix

In this section we study the eigenvalues and eigenvectors of H_L (see Remark 1.6) near an energy E' that is an eigenvalue of both H_0^+ and H_k^- (see the ends of Sections 4A3 and 4A4). We keep the notations of Sections 4A3 and 4A4.

Let $\varphi^+\in\ell^2(\mathbb{N})$ (resp. $\varphi^-\in\ell^2(\mathbb{Z}_-)$) be normalized eigenvectors of H_0^+ (resp. H_k^-) associated to E_-. Thus, by (4-28) and (4-32), we can pick, for $n\geq 0$ and $l\in\{0,\dots,p-1\}$,

$$\varphi^+_{np+l}=ca_l(E')\rho^n(E')\quad\text{and}\quad\varphi^-_{-np-l}=c^-b_l(E')\rho^n(E'). \tag{A-1}$$

Assume $L=Np+k$ and, for $l\in\{0,\dots,L\}$, define $\varphi^{\pm,L}\in\ell^2([\![0,L]\!])$ by

$$\varphi_l^{+,L}:=\varphi_l^+,\quad\varphi_{-1}^{+,L}=\varphi_{L+1}^{+,L}:=\varphi_{-1}^+=0,\quad\varphi_l^{-,L}:=\varphi_{l-L}^-\quad\text{and}\quad\varphi_{-1}^{-,L}=\varphi_{L+1}^{-,L}:=\varphi_0^-=0. \tag{A-2}$$

Thus, one has

$$H_L\varphi^{+,L}=E'\varphi^{+,L}+\varphi^{+}_{L+1}\delta_L,\quad H_L\varphi^{-,L}=E'\varphi^{-,L}+\varphi^{-}_{-L-1}\delta_0\quad\text{and}\quad\langle\varphi^{+,L},\varphi^{-,L}\rangle=O(N\rho^N(E)).\tag{A-3}$$

Recall that $a_k(E')\neq 0\neq b_k(E')$ (see Sections 4A3 and 4A4); thus, by (A-1), one has

$$|\varphi^{-}_{-L-1}|\asymp|\rho(E')|^n\asymp|\varphi^{+}_{L+1}|.\tag{A-4}$$

Moreover, as H_L converges to H_0^+ in the strong resolvent sense, for $\varepsilon>0$ sufficiently small and L sufficiently large, H_L has no spectrum in the compact $E'+\big[-2\varepsilon,-\frac{1}{2}\varepsilon\big]\cup\big[\frac{1}{2}\varepsilon,2\varepsilon\big]$. Let Π_L be the spectral projector onto the interval $\big[-\frac{1}{2}\varepsilon,\frac{1}{2}\varepsilon\big]$, that is, $\Pi_L:=1/(2i\pi)\int_{|z-E'|=\varepsilon}(H_L-z)^{-1}\,dz$. By (A-3), one computes

$$(1-\Pi_L)\varphi^{+,L}=\frac{\varphi^{+}_{L+1}}{2i\pi}\int_{|z-E'|=\varepsilon}(E'-z)^{-1}(H_L-z)^{-1}\delta_0\,dz.$$

Thus, one gets

$$\|(1-\Pi_L)\varphi^{+,L}\|+\|(1-\Pi_L)\varphi^{-,L}\|\lesssim|\rho(E')|^N.\tag{A-5}$$

Define

$$\tilde\chi^{+,L}=\frac{1}{\|\Pi_L\varphi^{+,L}\|}\Pi_L\varphi^{+,L}\quad\text{and}\quad\tilde\chi^{-,L}=\frac{1}{\|\Pi_L\varphi^{-,L}\|}\Pi_L\varphi^{-,L}.$$

The Gram matrix of $(\tilde\chi^{+,L},\tilde\chi^{-,L})$ then reads $\mathrm{Id}+O(N\rho^N(E))$. Orthonormalizing $(\tilde\chi^{+,L},\tilde\chi^{-,L})$ into $(\chi^{+,L},\chi^{-,L})$ and computing the matrix elements of $\Pi_L(H_L-E')$ in this basis, we obtain

$$\begin{pmatrix}\varphi^{+}_{L+1}\langle\delta_L,\varphi^{+,L}\rangle & \varphi^{+}_{L+1}\langle\delta_0,\varphi^{+,L}\rangle\\ \varphi^{-}_{-L-1}\langle\delta_L,\varphi^{-,L}\rangle & \varphi^{-}_{-L-1}\langle\delta_0,\varphi^{-,L}\rangle\end{pmatrix}+O(N^2\rho^{2N}(E))=\alpha\rho^N(E)\begin{pmatrix}0&1\\1&0\end{pmatrix}+O(N^2\rho^{2N}(E))$$

Thus, we obtain that the eigenvalues of H_L near E' are given by $E'\pm\alpha\rho^N(E)+O(N^2\rho^{2N}(E))$ and the eigenvectors by $\frac{1}{\sqrt{2}}(\varphi^{+,L}\pm\varphi^{-,L})+O(\rho^N(E))$. In particular, their components at 0 and L are asymptotic to nonvanishing constants.

References

[Aizenman 1994] M. Aizenman, "Localization at weak disorder: some elementary bounds", *Rev. Math. Phys.* **6**:5A (1994), 1163–1182. MR 1301371 Zbl 0843.47039

[Aizenman and Molchanov 1993] M. Aizenman and S. Molchanov, "Localization at large disorder and at extreme energies: an elementary derivation", *Comm. Math. Phys.* **157**:2 (1993), 245–278. MR 1244867 Zbl 0782.60044

[Aizenman et al. 2001] M. Aizenman, J. H. Schenker, R. M. Friedrich, and D. Hundertmark, "Finite-volume fractional-moment criteria for Anderson localization", *Comm. Math. Phys.* **224**:1 (2001), 219–253. MR 1868998 Zbl 1038.82038

[Barra and Gaspard 1999] F. Barra and P. Gaspard, "Scattering in periodic systems: from resonances to band structure", *J. Phys. A* **32**:18 (1999), 3357–3375. MR 1687454 Zbl 0962.81072

[Bougerol and Lacroix 1985] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics **8**, Birkhäuser, Boston, MA, 1985. MR 886674 Zbl 0572.60001

[Brown et al. 2005] B. M. Brown, S. Naboko, and R. Weikard, "The inverse resonance problem for Jacobi operators", *Bull. London Math. Soc.* **37**:5 (2005), 727–737. MR 2164835 Zbl 1148.39305

[Carmona 1983] R. Carmona, "One-dimensional Schrödinger operators with random or deterministic potentials: new spectral types", *J. Funct. Anal.* **51**:2 (1983), 229–258. MR 701057 Zbl 0516.60069

[Carmona and Lacroix 1990] R. Carmona and J. Lacroix, *Spectral theory of random Schrödinger operators*, Birkhäuser, Boston, MA, 1990. MR 1102675 Zbl 0717.60074

[Combes et al. 2009] J.-M. Combes, F. Germinet, and A. Klein, “Generalized eigenvalue-counting estimates for the Anderson model”, *J. Stat. Phys.* **135**:2 (2009), 201–216. MR 2505733 Zbl 1168.82016

[Comtet and Texier 1997] A. Comtet and C. Texier, “On the distribution of the Wigner time delay in one-dimensional disordered systems”, *J. Phys. A* **30**:23 (1997), 8017–8025. MR 1618595 Zbl 0937.82046

[Cycon et al. 1987] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Study ed., Springer, Berlin, 1987. MR 883643 Zbl 0619.47005

[Faris and Tsay 1989] W. G. Faris and W. J. Tsay, “Scattering of a wave packet by an interval of random medium”, *J. Math. Phys.* **30**:12 (1989), 2900–2903. MR 1025232 Zbl 0696.35127

[Faris and Tsay 1994] W. G. Faris and W. J. Tsay, “Time delay in random scattering”, *SIAM J. Appl. Math.* **54**:2 (1994), 443–455. MR 1265236 Zbl 0794.34074

[Germinet and Klopp 2011] F. Germinet and F. Klopp, “Spectral statistics for the discrete Anderson model in the localized regime”, pp. 11–24 in *Spectra of random operators and related topics* (Kyoto, 2009), edited by N. Minami, RIMS Kôkyûroku Bessatsu **B27**, Res. Inst. Math. Sci., Kyoto, 2011. MR 2885251 Zbl 1234.81075

[Germinet and Klopp 2014] F. Germinet and F. Klopp, “Spectral statistics for random Schrödinger operators in the localized regime”, *J. Eur. Math. Soc. (JEMS)* **16**:9 (2014), 1967–2031. MR 3273314 Zbl 06377398

[Helffer and Sjöstrand 1990] B. Helffer and J. Sjöstrand, “On diamagnetism and de Haas–van Alphen effect”, *Ann. Inst. H. Poincaré Phys. Théor.* **52**:4 (1990), 303–375. MR 1062904 Zbl 0715.35070

[Iantchenko and Korotyaev 2012] A. Iantchenko and E. Korotyaev, “Resonances for periodic Jacobi operators with finitely supported perturbations”, *J. Math. Anal. Appl.* **388**:2 (2012), 1239–1253. MR 2869822 Zbl 1250.47036

[Killip and Nakano 2007] R. Killip and F. Nakano, “Eigenfunction statistics in the localized Anderson model”, *Ann. Henri Poincaré* **8**:1 (2007), 27–36. MR 2299191 Zbl 1139.82022

[King 2009] F. W. King, *Hilbert transforms, I*, Encyclopedia of Mathematics and its Applications **124**, Cambridge University Press, 2009. MR 2542214 Zbl 1188.44005

[Kirsch 2008] W. Kirsch, “An invitation to random Schrödinger operators”, pp. 1–119 in *Random Schrödinger operators*, Panor. Synthèses **25**, Soc. Math. France, Paris, 2008. MR 2509110 Zbl 1162.82004

[Klopp 1995] F. Klopp, “An asymptotic expansion for the density of states of a random Schrödinger operator with Bernoulli disorder”, *Random Oper. Stochastic Equations* **3**:4 (1995), 315–331. MR 1373138 Zbl 0841.60047

[Klopp 2011] F. Klopp, “Decorrelation estimates for the eigenlevels of the discrete Anderson model in the localized regime”, *Comm. Math. Phys.* **303**:1 (2011), 233–260. MR 2775121 Zbl 1217.82091

[Klopp 2012] F. Klopp, “Resonances for “large” ergodic systems in one dimension: a review”, pp. 171–182 in *Spectral analysis of quantum Hamiltonians*, edited by R. Benguria et al., Oper. Theory Adv. Appl. **224**, Birkhäuser, Basel, 2012. MR 2962860 Zbl 1325.47071

[Klopp 2013] F. Klopp, “Asymptotic ergodicity of the eigenvalues of random operators in the localized phase”, *Probab. Theory Related Fields* **155**:3-4 (2013), 867–909. MR 3034795 Zbl 1281.82014

[Klopp $\geq$ 2016] F. Klopp, “Local Lyapunov behavior for localized one dimensional random Schrödinger operators”, in preparation.

[Kottos 2005] T. Kottos, “Statistics of resonances and delay times in random media: beyond random matrix theory”, *J. Phys. A* **38**:49 (2005), 10761–10786. MR 2197681 Zbl 1161.82342

[Kunz and Shapiro 2006] H. Kunz and B. Shapiro, “Resonances in a one-dimensional disordered chain”, *J. Phys. A* **39**:32 (2006), 10155–10160. MR 2252707 Zbl 1118.82025

[Kunz and Shapiro 2008] H. Kunz and B. Shapiro, “Statistics of resonances in a semi-infinite disordered chain”, *Phys. Rev. B* **77** (2008), art. ID 054203.

[Kunz and Souillard 1980] H. Kunz and B. Souillard, “Random Schrödinger operators. Some rigorous results”, pp. 83–87 in *Mathematical problems in theoretical physics* (Lausanne, 1979), edited by K. Osterwalder, Lecture Notes in Phys. **116**, Springer, Berlin, 1980. MR 582611 Zbl 0446.60049

[Lifshits et al. 1988] I. M. Lifshits, S. A. Gredeskul, and L. A. Pastur, *Introduction to the theory of disordered systems*, John Wiley & Sons, New York, 1988. MR 1042095

[Mather 1971] J. N. Mather, "On Nirenberg's proof of Malgrange's preparation theorem", pp. 116–120 in *Proceedings of Liverpool Singularities—Symposium, I (1969/70)*, Lecture Notes in Mathematics **192**, Springer, Berlin, 1971. MR 0412461 Zbl 0211.56102

[Minami 1996] N. Minami, "Local fluctuation of the spectrum of a multidimensional Anderson tight binding model", *Comm. Math. Phys.* **177**:3 (1996), 709–725. MR 1385082 Zbl 0851.60100

[van Moerbeke 1976] P. van Moerbeke, "The spectrum of Jacobi matrices", *Invent. Math.* **37**:1 (1976), 45–81. MR 0650253 Zbl 0361.15010

[Molchanov 1982] S. A. Molchanov, "The local structure of the spectrum of a random one-dimensional Schrödinger operator", *Trudy Sem. Petrovsk.* **8** (1982), 195–210. In Russian; translated in *J. Soviet Math.* **32**:4 (1986), 371–383. MR 673165 Zbl 0509.60064

[Pastur and Figotin 1992] L. Pastur and A. Figotin, *Spectra of random and almost-periodic operators*, Grundlehren der Math. Wissenschaften **297**, Springer, Berlin, 1992. MR 1223779 Zbl 0752.47002

[Pavlov 1994] B. Pavlov, "Nonphysical sheet for perturbed Jacobian matrices", *Algebra i Analiz* **6**:3 (1994), 185–199. Reprinted in *St. Petersburg J. Math* **6**:3 (1995), 619–633. MR 1301837 Zbl 0829.47011

[Reed and Simon 1980] M. Reed and B. Simon, *Methods of modern mathematical physics, I: Functional analysis*, 2nd ed., Academic Press, New York, 1980. MR 751959

[Simon 1995] B. Simon, "Spectral analysis of rank one perturbations and applications", pp. 109–149 in *Mathematical quantum theory, II: Schrödinger operators* (Vancouver, 1993), edited by J. Feldman et al., CRM Proc. Lecture Notes **8**, Amer. Math. Soc., Providence, RI, 1995. MR 1332038 Zbl 0824.47019

[Simon 2005] B. Simon, *Trace ideals and their applications*, 2nd ed., Mathematical Surveys and Monographs **120**, Amer. Math. Soc., Providence, RI, 2005. MR 2154153 Zbl 1074.47001

[Sjöstrand and Zworski 1991] J. Sjöstrand and M. Zworski, "Complex scaling and the distribution of scattering poles", *J. Amer. Math. Soc.* **4**:4 (1991), 729–769. MR 1115789 Zbl 0752.35046

[Teschl 2000] G. Teschl, *Jacobi operators and completely integrable nonlinear lattices*, Mathematical Surveys and Monographs **72**, Amer. Math. Soc., Providence, RI, 2000. MR 1711536 Zbl 1056.39029

[Texier and Comtet 1999] C. Texier and A. Comtet, "Universality of the Wigner time delay distribution for one-dimensional random potentials", *Phys. Rev. Lett.* **82** (1999), 4220–4223.

[Titov and Fyodorov 2000] M. Titov and Y. V. Fyodorov, "Time-delay correlations and resonances in one-dimensional disordered systems", *Phys. Rev. B* **61** (2000), R2444–R2447.

[Zworski 2002] M. Zworski, "Quantum resonances and partial differential equations", pp. 243–252 in *Proceedings of the International Congress of Mathematicians, III* (Beijing, 2002), edited by T. Li, Higher Ed. Press, Beijing, 2002. MR 1957536 Zbl 1136.35432

Received 2 Oct 2012. Revised 29 Aug 2015. Accepted 30 Jan 2016.

FRÉDÉRIC KLOPP: frederic.klopp@imj-prg.fr
Institute de Mathématiques Jussieu — Paris Rive Gauche, UMR 7586, CNRS, Université Pierre et Marie Curie, Case 186, 4 place Jussieu, 75252 Paris, France

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.353

ON CHARACTERIZATION OF THE SHARP STRICHARTZ INEQUALITY FOR THE SCHRÖDINGER EQUATION

JIN-CHENG JIANG AND SHUANGLIN SHAO

We study the extremal problem for the Strichartz inequality for the Schrödinger equation on $\mathbb{R} \times \mathbb{R}^2$. We show that the solutions to the associated Euler–Lagrange equation are exponentially decaying in the Fourier space and thus can be extended to be complex analytic. Consequently, we provide a new proof of the characterization of the extremal functions: the only extremals are Gaussian functions, as investigated previously by Foschi, Hundertmark and Zharnitsky.

1. Introduction

We begin with some notation. For a Schwarz function f on $\mathbb{R}^d$, $d \geq 1$, define the Fourier transform

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx, \quad \xi \in \mathbb{R}^d.$$

The inverse of the Fourier transform,

$$\mathcal{F}^{-1}(f)(x) = f^{\vee}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} f(\xi)\, d\xi, \quad x \in \mathbb{R}^d.$$

The linear Strichartz inequality for the Schrödinger equation [Keel and Tao 1998; Tao 2006] asserts that

$$\|e^{it\Delta} f\|_{L^{2+4/d}_{t,x}(\mathbb{R}\times\mathbb{R}^d)} \leq C_d \|f\|_{L^2(\mathbb{R}^d)}, \tag{1}$$

where $e^{it\Delta} f(x) = (1/(2\pi)^d) \int_{\mathbb{R}^d} e^{ix\cdot\xi+it|\xi|^2} \hat{f}(\xi)\, d\xi$. We specify $d = 2$ and consider

$$\|e^{it\Delta} f\|_{L^4_{t,x}(\mathbb{R}\times\mathbb{R}^2)} \leq \boldsymbol{R} \|f\|_{L^2(\mathbb{R}^2)}, \tag{2}$$

where

$$\boldsymbol{R} := \sup\left\{ \frac{\|e^{it\Delta} f\|_{L^4_{t,x}(\mathbb{R}\times\mathbb{R}^2)}}{\|f\|_{L^2(\mathbb{R}^2)}} : f \in L^2,\ f \neq 0 \right\}. \tag{3}$$

We define an extremal function or extremal to (2) to be a nonzero function $f \in L^2$ such that the inequality is optimized, in the sense that

$$\|e^{it\Delta} f\|_{L^4_{t,x}(\mathbb{R}\times\mathbb{R}^2)} = \boldsymbol{R} \|f\|_{L^2(\mathbb{R}^2)}. \tag{4}$$

The extremal problem of (2) concerns:

(i) Whether there exists an extremal function?

MSC2010: 35J10.
Keywords: Schrödinger equation, Strichartz inequality and extremals.

(ii) How to characterize the extremal functions? What are the explicit forms of extremal functions? Are they unique up to the symmetry of the inequality?

From Foschi [2007] and Hundertmark and Zharnitsky [2006], it is known that the Gaussian functions are the only extremal functions of the linear Strichartz inequality (2) for the dimensions $d = 1, 2$. Here Gaussian functions $\mathbb{R}^d \to \mathbb{C}$, $d = 1, 2$, are of the form

$$e^{A|x|^2+B\cdot x+C}$$

with $A, C \in \mathbb{C}$, $B \in \mathbb{C}^d$ and the real part of A negative. The existence of extremizers was established previously by Kunze [2003] for the Strichartz inequality (1) when $d = 1$. When $d \geq 3$, existence of extremizers is proved by the second author in [Shao 2009] .

In this note, we are interested in the problem of how to characterize extremals for (2) via the study of the associated Euler–Lagrange equation. We show that the solutions of this generalized Euler–Lagrange equation enjoy fast decay in the Fourier space and thus can be extended to be complex analytic; see Theorem 1.1. Then, as an easy consequence, we give an alternative proof that all extremal functions to (2) are Gaussians, based on solving a functional equation of extremizers derived in [Foschi 2007]; see (7) and Theorem 1.2. Indeed, in the proof given below we use the information that f is twice continuously differentiable, i.e., $f \in C^2$, which can be lowered to continuity by a more refined argument. The functional inequality (7) is a key ingredient in Foschi's proof. To prove f in (7) to be a Gaussian function, local integrability of f is assumed in [Foschi 2007], which is further reduced to measurable functions in [Charalambides 2013].

Let f be an extremal function to (2) with the constant $\boldsymbol{R}$. Then f satisfies the generalized Euler–Lagrange equation

$$\omega\langle g, f\rangle = \mathfrak{Q}(g, f, f, f) \quad \text{for all } g \in L^2, \tag{5}$$

where $\omega = \mathfrak{Q}(f, f, f, f)/\|f\|_{L^2}^2 > 0$ and $\mathfrak{Q}(f_1, f_2, f_3, f_4)$ is the integral

$$\int_{(\mathbb{R}^2)^4} \bar{\hat{f}}_1(\xi_1)\bar{\hat{f}}_2(\xi_2)\hat{f}_3(\xi_3)\hat{f}_4(\xi_4)\delta(\xi_1+\xi_2-\xi_3-\xi_4)\delta(|\xi_1|^2+|\xi_2|^2-|\xi_3|^2-|\xi_4|^2)\,d\xi_1\,d\xi_2\,d\xi_3\,d\xi_4 \tag{6}$$

for $f_i \in L^2(\mathbb{R}^2)$, $1 \leq i \leq 4$, and $\delta(\xi) = (2\pi)^{-d}\int_{\mathbb{R}^d} e^{i\xi\cdot x}\,dx$ in the distribution sense for $d = 1, 2$. The proof of (5) is standard; see, e.g., [Evans 2010, p. 489] or [Hundertmark and Lee 2012, Section 2] for similar derivations of Euler–Lagrange equations.

Theorem 1.1. *If f solves the generalized Euler–Lagrange equation* (5) *for some $\omega > 0$, then there exists $\mu > 0$ such that*

$$e^{\mu|\xi|^2}\hat{f} \in L^2(\mathbb{R}^2).$$

Furthermore, f can be extended to be complex analytic on $\mathbb{C}^2$.

To prove this theorem, we follow the argument in [Hundertmark and Shao 2012]. Similar reasoning has appeared previously in [Erdoğan et al. 2011; Hundertmark and Lee 2009]. It relies on a multilinear weighted Strichartz estimate and a continuity argument. See Lemmas 2.1 and 2.2.

Next we prove that the extremals to (2) are Gaussian functions. We start with the study of the functional equation derived in [Foschi 2007], which reads

$$f(x)f(y) = f(w)f(z) \tag{7}$$

for any $x, y, w, z \in \mathbb{R}^2$ such that

$$x + y = w + z \quad \text{and} \quad |x|^2 + |y|^2 = |w|^2 + |z|^2. \tag{8}$$

Note that $x, y, w, z \in \mathbb{R}^2$ satisfy the relation (8) if and only if these four points form a rectangle in $\mathbb{R}^2$ with vertices x, y, w and z. Indeed, by (8), these four points x, y, w and z form a parallelogram on $\mathbb{R}^2$ and $x \cdot y = w \cdot z$. Secondly, $w - x$ is perpendicular to $z - x$, since $(w-x)\cdot(z-x) = w\cdot z - w\cdot x - x\cdot z + |x|^2 = w \cdot z - (x + y)\cdot x + |x|^2 = w \cdot z - y \cdot x = 0$. This proves that x, y, w and z form a rectangle on $\mathbb{R}^2$. In [Foschi 2007], it is proven that $f \in L^2$ satisfies (7) if and only if f is an extremal function to (2). Basically, this comes from two aspects. One is that, in the Foschi's proof of the sharp Strichartz inequality, only the Cauchy–Schwarz inequality is used at one place besides equality. So the equality in the Strichartz inequality (2), or equivalently the equality in Cauchy-Schwarz, yields the same functional equation as (7), where f is replaced by $\hat{f}$. The other one is that the Strichartz norm for the Schrödinger equation satisfies the identity

$$\|e^{it\Delta} f\|_{L^4(\mathbb{R}\times\mathbb{R}^2)} = C\|e^{it\Delta} f^\vee\|_{L^4(\mathbb{R}\times\mathbb{R}^2)} \tag{9}$$

for some $C > 0$.

Foschi [2007] is able to show that all the solutions to (7) are Gaussians under the assumption that f is a locally integrable function. This can be viewed as an investigation of the Cauchy functional equation (7) for functions supported on the paraboloids. To characterize the extremals for the Tomas–Stein inequality for the sphere in $\mathbb{R}^3$, [Christ and Shao 2012] studies the same functional equation (7) for functions supported on the sphere and prove that they are exponentially affine functions. Charalambides [2013] generalizes the analysis in [Christ and Shao 2012] to some general hypersurfaces in $\mathbb{R}^n$ that include the sphere, paraboloids and cones as special examples and proves that the solutions are exponentially affine functions. In [Charalambides 2013; Christ and Shao 2012], the functions are assumed to be measurable functions.

By the analyticity established in Theorem 1.1, equations (7) and (8) have the following easy consequence, which recovers the result in [Foschi 2007; Hundertmark and Zharnitsky 2006].

Theorem 1.2. *Suppose that f is an extremal function to* (2). *Then*

$$f(x) = e^{A|x|^2 + B\cdot x + C}, \tag{10}$$

where $A, C \in \mathbb{C}$, $B \in \mathbb{C}^2$ and $\Re(A) < 0$.

Let f be an extremal function to (2). Then, by Theorem 1.1, f is continuous. This, together with (7) and (8), implies that any nontrivial f is nowhere vanishing on $\mathbb{R}^2$; see, e.g., [Foschi 2007, Lemma 7.13]. For any $a \in \mathbb{R}^2$, there is a disk $D(a, r) \subset \mathbb{C}^2$, $r > 0$, such that f is C^2 by Theorem 1.1 and f is nowhere vanishing. Then $\log f$ is C^2 on $D(a, r)$; see, e.g., [Krantz 1992, Lemma 6.1.9]. Similar claims can be made for $\log f^2$. Then, up to a multiple of 2π,

$$\log f^2(a) = \log f(a) + \log f(a).$$

After restriction to $\mathbb{R}^2$, f satisfies (7) for x, y, w and z satisfying (8). So, by taking r sufficiently small,

$$\log f(x) + \log f(y) = \log f(w) + \log f(z)$$

for $x, y, w, z \in B(a, r) \subset \mathbb{R}^2$ related as in (8). Since $\log f$ is twice differentiable, it is not hard to see that $\log f$ is a quadratic polynomial on $B(a, r)$. So $\log f$ is a quadratic polynomial on $\mathbb{R}^2$. Indeed, let $a = 0$ and $\phi(x_1) = \log f(x_1, 0)$, $\psi(0, x_2) = \log f(0, x_2)$. Then, since the four points (x_1, x_2), $(x_2, -x_1)$, $(x_1 + x_2, x_2 - x_1)$ and $(0, 0)$ satisfy (8), we see that

$$[\phi(x_1) + \psi(x_2)] + [\phi(x_2) + \psi(-x_1)] = [\phi(x_1 + x_2) + \psi(x_2 - x_1)] + \log f(0, 0).$$

By differentiating firstly in x_1 and then in x_2, we see that $\phi'' = \psi''$ is a constant. Thus f is a quadratic polynomial. It is easy to see that this argument generalizes to any $a \in \mathbb{R}^2$.

2. Complex analyticity

In this section, we show that the solutions to the generalized Euler–Lagrange equation (5) can be extended to be complex analytic.

We define

$$\begin{aligned} \eta &:= (\eta_1, \eta_2, \eta_3, \eta_4) \in (\mathbb{R}^2)^4, \\ a(\eta) &:= \eta_1 + \eta_2 - \eta_3 - \eta_4, \\ b(\eta) &:= |\eta_1|^2 + |\eta_2|^2 - |\eta_3|^2 - |\eta_4|^2. \end{aligned}$$

Let $\varepsilon \geq 0$ and $\mu \geq 0$. For $\xi \in \mathbb{R}^2$, define

$$F(\xi) := F_{\mu,\varepsilon}(\xi) = \frac{\mu|\xi|^2}{1 + \varepsilon|\xi|^2}. \tag{11}$$

Define the weighted multilinear integral for $h_i \in L^2(\mathbb{R}^2)$, $1 \leq i \leq 4$, by

$$M_F(h_1, h_2, h_3, h_4) := \int_{(\mathbb{R}^2)^4} e^{F(\eta_1) - \sum_{j=2}^4 F(\eta_j)} \prod_{j=1}^4 |h(\eta_j)| \delta(a(\eta)) \delta(b(\eta))\, d\eta. \tag{12}$$

The multilinear estimate we need shows the weak interaction of Schrödinger waves between the high and low frequency. More precisely:

Lemma 2.1. *Let $h_i \in L^2(\mathbb{R}^2)$, $1 \leq i \leq 4$, and let $s > 1$ be a large number. If the Fourier transforms of h_1 and h_2 are supported in $\{\xi : |\xi| \leq s\}$ and $\{\xi : |\xi| \geq Ns\}$ with $N > 1$ a large number, respectively, then*

$$M_F(h_1, h_2, h_3, h_4) \leq CN^{-1/2} \prod_{j=1}^4 \|h_j\|_{L^2}. \tag{13}$$

Proof. The proof of this lemma needs the following two inequalities:

$$M_F(h_1, h_2, h_3, h_4) \le \int_{(\mathbb{R}^2)^4} \prod_{j=1}^{4} |h_j(\eta_j)| \delta(a(\eta)) \delta(b(\eta))\, d\eta \tag{14}$$

and

$$\|e^{it\Delta} h_1 e^{it\Delta} h_2\|_{L^2_{t,x}} \le C N^{-1/2} \|h_1\|_{L^2} \|h_2\|_{L^2}. \tag{15}$$

Together with the Cauchy–Schwarz inequality and the $L^2 \to L^4$ Strichartz inequality, the inequality (13) follows from (14) and (15). Note that (15) is established in [Bourgain 1998]. Thus it remains to establish (14), where we follow [Erdoğan et al. 2011; Hundertmark and Shao 2012].

On the support of η determined by $\delta(a(\eta))$ and $\delta(b(\eta))$, we have

$$\eta_1 + \eta_2 = \eta_3 + \eta_4 \quad \text{and} \quad |\eta_1|^2 + |\eta_2|^2 = |\eta_3|^2 + |\eta_4|^2.$$

Thus,

$$|\eta_1|^2 \le |\eta_2|^2 + |\eta_3|^2 + |\eta_4|^2.$$

Since the function $x \mapsto x/(1+\varepsilon x)$ is increasing on the interval $[0, \infty)$, we have

$$\frac{|\eta_1|^2}{1+\varepsilon|\eta_1|^2} \le \frac{\sum_{j=2}^4 |\eta_j|^2}{1+\sum_{j=2}^4 \varepsilon|\eta_j|^2} = \sum_{j=2}^4 \frac{|\eta_j|^2}{1+\sum_{j=2}^4 \varepsilon|\eta_j|^2} \le \sum_{j=2}^4 \frac{|\eta_j|^2}{1+\varepsilon|\eta_j|^2}.$$

This implies that $F(\eta_1) \le \sum_{j=2}^4 F(\eta_j)$, since $\mu \ge 0$. Hence,

$$e^{F(\eta_1) - \sum_{j=2}^4 F(\eta_j)} \le 1.$$

Therefore (14) follows by taking the absolute value in the integral. □

If $f \in L^2$ satisfies the generalized Euler–Lagrange equation (5), the following bootstrap lemma shows that f gains certain regularity; namely, there is a constant $\mu > 0$ depending on the function f such that $e^{\mu|\xi|^2} \hat{f} \in L^2$. This is enough to conclude that f can be extended to be complex analytic.

Lemma 2.2. *If f solves the generalized Euler–Lagrange equation* (5) *for some $\omega > 0$ and $\|f\|_{L^2} = 1$, then for $\hat{f}_> := \hat{f} 1_{|\xi| \ge s^2}$ with $s > 0$, there is a large constant $s \gg 1$ such that, for $\mu = s^{-4}$,*

$$\omega \|e^{F(\cdot)} \hat{f}_>\|_{L^2} \le o_1(1) \|e^{F(\cdot)} \hat{f}_>\|_{L^2} + C\|e^{F(\cdot)} \hat{f}_>\|_{L^2}^2 + C\|e^{F(\cdot)} \hat{f}_>\|_{L^2}^3 + o_2(1), \tag{16}$$

where $\lim_{s\to\infty} o_i(1) = 0$ uniformly for all $\varepsilon > 0$, $i = 1, 2$, and the constant $C > 0$ is independent of ε and s.

Proof. Define $h(\xi) = e^{F(\xi)} \hat{f}(\xi)$ and $h_>(\xi) = e^{F(\xi)} \hat{f}_>$, where $\hat{f}_> = \hat{f} 1_{|\xi| \ge s^2}$. Let P denote the symbol of differentiation $-i\partial_x$; under the Fourier transform, $\widehat{P} = |\xi|$. Correspondingly, we write $F(P)$ with the Fourier symbol $\mu|\xi|^2/(1+\varepsilon|\xi|^2)$.

We expand

$$\|e^{F(\cdot)} \hat{f}_>\|_{L^2}^2 = \langle e^{F(\cdot)} \hat{f}_>, e^{F(\cdot)} \hat{f}_> \rangle = \langle e^{2F(\cdot)} \hat{f}_>, \hat{f} \rangle = \langle e^{2F(P)} f_>, f \rangle.$$

Thus, in the generalized Euler–Lagrange equation (5), setting $g = e^{2F(P)} f_>$, we see that

$$\omega \| e^{F(P)} f_> \|_{L^2}^2 = Q(e^{2F(P)} f_>, f, f, f). \tag{17}$$

Since $\hat{f} = e^{-F(\xi)} h$ and $e^{2F(\xi)} \hat{f}_> = e^{F(\xi)} h_>$,

$$\begin{aligned} Q(e^{2F(P)} f_>, f, f, f) &= \int_{(\mathbb{R}^2)^4} e^{2F(\xi_1)} \overline{\hat{f}}_>(\xi_1) \overline{\hat{f}}_>(\xi_2) \hat{f}(\xi_3) \hat{f}_4(\xi_4) \delta(a(\xi)) \delta(b(\xi))\, d\xi \\ &= \int_{(\mathbb{R}^2)^4} \overline{e^{F(\xi_1)} h_>(\xi_1)}\, \overline{e^{-F(\xi_2)} h(\xi_2)} e^{-F(\xi_3)} h(\xi_3) e^{-F(\xi_4)} h(\xi_4) \delta(a(\xi)) \delta(b(\xi))\, d\xi \\ &= \int_{(\mathbb{R}^2)^4} e^{F(\xi_1) - \sum_{j=2}^4 F(\xi_j)} h_>(\xi_1) h(\xi_2) h(\xi_3) h(\xi_4) \delta(a(\xi)) \delta(b(\xi))\, d\xi, \end{aligned}$$

where $a(\xi) = \xi_1 + \xi_2 - \xi_3 - \xi_4$ and $b(\xi) = |\xi_1|^2 + |\xi_2|^2 - |\xi_3|^2 - |\xi_4|^2$ for $\xi = (\xi_1, \xi_2, \xi_3, \xi_4) \in (\mathbb{R}^2)^4$. Thus,

$$\omega \| e^{F(P)} f_> \|_{L^2}^2 \le M_F(h_>, h, h, h). \tag{18}$$

Define

$$h_\sim = h 1_{s \le |\xi| \le s^2}, \; h_\ll = h 1_{|\xi| < s} \quad \text{and} \quad h_< = h_\ll + h_\sim .$$

We split the integral $M_F(h_>, h, h, h)$ into the following pieces:

$$M_F(h_>, h_<, h_<, h_<) + \sum_{j_2, j_3, j_4} M_F(h_>, h_{j_2}, h_{j_3}, h_{j_4}) =: A + B,$$

where h_{j_k} is either $h_>$ or $h_<$, but at least one is $h_>$. We further split A into two terms,

$$M_F(h_>, h_\ll, h_<, h_<) + M_F(h_>, h_\sim, h_<, h_<);$$

we estimate this term by using Lemma 2.1:

$$A \lesssim s^{-1/2} \|h_>\|_{L^2} \|h_\ll\|_{L^2} \|h_<\|_{L^2}^2 + \|h_>\|_{L^2} \|h_\sim\|_{L^2} \|h_<\|_{L^2}^2 \lesssim \|h_>\|_{L^2} (s^{-1/2} \|h_\ll\|_{L^2} + \|h_\sim\|_{L^2}) \|h_<\|_{L^2}^2.$$

Since $\|f\|_{L^2} = 1$,

$$\begin{aligned} \|h_<\|_{L^2} &\le e^{\mu s^4} \|f\|_{L^2} = e^{\mu s^4}, \\ \|h_\ll\|_{L^2} &\le e^{\mu s^2}, \\ \|h_\sim\|_{L^2} &\le e^{\mu s^4} \|f_\sim\|_{L^2}, \end{aligned}$$

where $f_\sim$ is defined by $\hat{f}_\sim = \hat{f} 1_{s \le |\xi| \le s^2}$. Thus we have

$$A \lesssim e^{3\mu s^4} \|h_>\|_{L^2} (s^{-1/2} e^{\mu s^2 - \mu s^4} + \|f_\sim\|_{L^2}). \tag{19}$$

Similarly we estimate the term B. We split B as $B_1 + B_2$, where $B_1 = \sum_{j_2, j_3, j_4} M_F(h_>, h_{j_2}, h_{j_3}, h_{j_4})$ contains exactly one $h_>$ in $\{h_{j_2}, h_{j_3}, h_{j_4}\}$, while $B_2 = \sum_{j_2, j_3, j_4} M_F(h_>, h_{j_2}, h_{j_3}, h_{j_4})$ contains two or more $h_>$.

To estimate B_1,

$$B_1 \lesssim e^{\mu s^4}\|h_>\|_{L^2}^2\|h_<\|_{L^2}(s^{-1/2}e^{\mu s^2-\mu s^4}+\|f_\sim\|_{L^2}) \lesssim e^{2\mu s^4}\|h_>\|_{L^2}^2(s^{-1/2}e^{\mu s^2-\mu s^4}+\|f_\sim\|_{L^2}). \quad (20)$$

To estimate B_2,

$$B_2 \lesssim \|h_>\|_{L^2}^3\|h_<\|_{L^2}+\|h_>\|_{L^2}^4 \lesssim e^{\mu s^4}\|h_>\|_{L^2}^3+\|h_>\|_{L^2}^4. \quad (21)$$

Thus, from (19), (20) and (21), we obtain

$$\|e^{F(\cdot)}\hat{f}_>\|_{L^2}^2$$
$$\lesssim e^{3\mu s^4}\|h_>\|_{L^2}(s^{-1/2}e^{\mu s^2-\mu s^4}+\|f_\sim\|_{L^2})+e^{2\mu s^4}\|h_>\|_{L^2}^2(s^{-1/2}e^{\mu s^2-\mu s^4}+\|f_\sim\|_{L^2})+e^{\mu s^4}\|h_>\|_{L^2}^3+\|h_>\|_{L^2}^4.$$

Since $\lim_{s\to\infty}\|f_\sim\|_{L^2}=0$, we take s sufficiently large and set $\mu=s^{-4}$:

$$\omega\|e^{F(\cdot)}\hat{f}_>\|_{L^2} \le o_1(1)\|e^{F(\cdot)}\hat{f}_>\|_{L^2}+C\|e^{F(\cdot)}\hat{f}_>\|_{L^2}^2+C\|e^{F(\cdot)}\hat{f}_>\|_{L^2}^3+o_2(1), \quad (22)$$

which completes the proof of Lemma 2.2. □

Remark 2.3. Clearly the choice of μ in the preceding lemma depends on the function f itself.

Now we conclude that f in Lemma 2.2 gains certain regularity.

Proof of Theorem 1.1. Let $f \in L^2$ and $f \ne 0$. We normalize f so that $\|f\|_{L^2}=1$. In Lemma 2.2, we choose s sufficiently large such that $o_1(1) \le \frac{1}{2}\omega$ and $o_2(1) \le \frac{1}{2}M$, where $M=\sup\{G(x) : x \in [0,\infty)\}$, and

$$G(x) := \tfrac{1}{2}\omega x - Cx^2 - Cx^3, \quad x \in [0,\infty), \quad (23)$$

and C is the same constant as in (16). It is easy to see that $0 \le M < \infty$. Then $G(x) \le M$ for all $x \in [0,\infty)$ by Lemma 2.2. Also the function G is continuous on $[0,\infty)$. On the other hand, $G''(x) < 0$ for all $x \in (0,\infty)$; thus G is concave. The line $G=\frac{1}{2}M$ intersects at two points of the positive x axis, $x=x_0$ and $x=x_1>0$.

We define $H : (0,\infty) \to [0,\infty)$ via

$$H(\varepsilon)=\left(\int_{|\xi|\ge s^2}|e^{F_{s^{-4},\varepsilon}(\xi)}\hat{f}|^2\,d\xi\right)^{\frac{1}{2}}.$$

The function H is continuous on $(0,\infty)$ by the dominated convergence theorem and $H(0,\infty)$ is connected. Hence $G^{-1}\big([0,\frac{1}{2}M]\big)$ is either contained in $[0,x_0]$ or $[x_1,\infty)$; only one alternative holds. For $\varepsilon=1$ and s sufficiently large, $H(1) \ge x_1$ is impossible. Hence the first alternative holds.

Therefore $G^{-1}\big([0,\frac{1}{2}M]\big) \subset [0,x_0]$, which yields that

$$\|e^{F(\cdot)}\hat{f}_>\|_{L^2} \le C_0, \quad \text{that is,} \quad \|e^{s^{-4}|\xi|^2/(1+\varepsilon|\xi|^2)}\hat{f}_>\|_{L^2} \le C_0, \quad (24)$$

uniformly in all $\varepsilon > 0$. By the monotone convergence theorem,

$$\|e^{s^{-4}|\xi|^2}\hat{f}_>\|_{L^2} \le C_0 < \infty.$$

It is clear that $e^{s^{-4}|\xi|^2}\hat{f}1_{|\xi|\le s^2}\in L^2$. Therefore,

$$e^{s^{-4}|\xi|^2}\hat{f}\in L^2.$$

Let $\mu=s^{-4}$. This proves the first half of Theorem 1.1.

To prove that f can be extended to be complex analytic on $\mathbb{C}^2$, we observe that, by the Cauchy–Schwarz inequality, for any $\lambda\in\mathbb{R}$,

$$e^{\lambda|\xi|}\hat{f}(\xi)=e^{\lambda|\xi|-\mu|\xi|^2}e^{\mu|\xi|^2}\hat{f}(\xi)\in L^2(\mathbb{R}^2). \tag{25}$$

So it is not hard to see that f can be extended to be complex analytic on $\mathbb{C}^2$; see, e.g., [Reed and Simon 1975, Theorem IX.13]. Alternatively, analyticity can be obtained in the following way. Similarly to in (25) for $k\in\mathbb{N}\cup\{0\}$, $|\xi|^k e^{\lambda|\xi|}\hat{f}\in L^1(\mathbb{R}^2)$. For $z\in\mathbb{C}^2$, we choose $\lambda>|z|$, then

$$f(z)=(2\pi)^{-2}\int_{\mathbb{R}^2}e^{iz\cdot\xi-\lambda|\xi|}e^{\lambda|\xi|}\hat{f}(\xi)\,d\xi.$$

Then, by taking differentiation under the integral sign, complex analyticity follows. □

Acknowledgements

The research of Jiang is supported by National Science Council Grant NSC 102-2115-M-007-0101-MY 2. Shao is supported in part by the NSF grant DMS-1160981. The authors would like to thank the anonymous referee for helpful comments and suggestions, which have been incorporated into this paper.

References

[Bourgain 1998] J. Bourgain, "Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity", *Internat. Math. Res. Notices* **1998**:5 (1998), 253–283. MR 99f:35184 Zbl 0917.35126

[Charalambides 2013] M. Charalambides, "On restricting Cauchy–Pexider functional equations to submanifolds", *Aequationes Math.* **86**:3 (2013), 231–253. MR 3127007 Zbl 1281.39021

[Christ and Shao 2012] M. Christ and S. Shao, "On the extremizers of an adjoint Fourier restriction inequality", *Adv. Math.* **230**:3 (2012), 957–977. MR 2921167 Zbl 1258.35007

[Erdoğan et al. 2011] M. B. Erdoğan, D. Hundertmark, and Y.-R. Lee, "Exponential decay of dispersion managed solitons for vanishing average dispersion", *Math. Res. Lett.* **18**:1 (2011), 11–24. MR 2012f:35501 Zbl 1239.35021

[Evans 2010] L. C. Evans, *Partial differential equations*, 2nd ed., Graduate Studies in Mathematics **19**, American Mathematical Society, Providence, RI, 2010. MR 2011c:35002 Zbl 1194.35001

[Foschi 2007] D. Foschi, "Maximizers for the Strichartz inequality", *J. Eur. Math. Soc.* **9**:4 (2007), 739–774. MR 2008k:35389 Zbl 1231.35028

[Hundertmark and Lee 2009] D. Hundertmark and Y.-R. Lee, "Decay estimates and smoothness for solutions of the dispersion managed non-linear Schrödinger equation", *Comm. Math. Phys.* **286**:3 (2009), 851–873. MR 2009k:35299 Zbl 1173.35686

[Hundertmark and Lee 2012] D. Hundertmark and Y.-R. Lee, "On non-local variational problems with lack of compactness related to non-linear optics", *J. Nonlinear Sci.* **22**:1 (2012), 1–38. MR 2012m:49003 Zbl 1244.49008

[Hundertmark and Shao 2012] D. Hundertmark and S. Shao, "Analyticity of extremizers to the Airy–Strichartz inequality", *Bull. Lond. Math. Soc.* **44**:2 (2012), 336–352. MR 2914611 Zbl 1280.35126

[Hundertmark and Zharnitsky 2006] D. Hundertmark and V. Zharnitsky, "On sharp Strichartz inequalities in low dimensions", *Int. Math. Res. Not.* **2006** (2006), Art. ID 34080, 18. MR 2007b:35277 Zbl 1131.35308

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR 2000d:35018 Zbl 0922.35028

[Krantz 1992] S. G. Krantz, *Function theory of several complex variables*, 2nd ed., Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992. MR 93c:32001 Zbl 0776.32001

[Kunze 2003] M. Kunze, "On the existence of a maximizer for the Strichartz inequality", *Comm. Math. Phys.* **243**:1 (2003), 137–162. MR 2004i:35006 Zbl 1060.35133

[Reed and Simon 1975] M. Reed and B. Simon, *Methods of modern mathematical physics, II: Fourier analysis, self-adjointness*, Academic Press, New York, 1975. MR 58 #12429b Zbl 0308.47002

[Shao 2009] S. Shao, "Maximizers for the Strichartz and the Sobolev–Strichartz inequalities for the Schrödinger equation", *Electron. J. Differential Equations* **2009** (2009), art. ID #3. MR 2010a:35032 Zbl 1173.35692

[Tao 2006] T. Tao, *Nonlinear dispersive equations: local and global analysis*, CBMS Regional Conference Series in Mathematics **106**, American Mathematical Society, Providence, RI, 2006. MR 2008i:35211 Zbl 1106.35001

Received 22 Apr 2014. Revised 28 Sep 2015. Accepted 16 Dec 2015.

JIN-CHENG JIANG: jcjiang@math.nthu.edu.tw
Department of Mathematics, National Tsing-Hua University, No. 101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan

SHUANGLIN SHAO: slshao@ku.edu
Department of Mathematics, University of Kansas, 615 Snow Hall, 1460 Jayhawk Blvd, Lawrence, KS 66045-7594, United States

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.363

FUTURE ASYMPTOTICS AND GEODESIC COMPLETENESS OF POLARIZED T^2-SYMMETRIC SPACETIMES

PHILIPPE G. LEFLOCH AND JACQUES SMULEVICI

We investigate the late-time asymptotics of future-expanding, polarized vacuum Einstein spacetimes with T^2-symmetry on T^3, which, by definition, admit two spacelike Killing fields. Our main result is the existence of a stable asymptotic regime within this class; that is, we provide here a full description of the late-time asymptotics of the solutions to the Einstein equations when the initial data set is close to the asymptotic regime. Our proof is based on several energy functionals with lower-order corrections (as is standard for such problems) and the derivation of a simplified model that we exhibit here. Roughly speaking, the Einstein equations in the symmetry class under consideration consist of a system of wave equations coupled to constraint equations plus a system of ordinary differential equations. The unknowns involved in the system of ordinary equations are blowing up in the future timelike directions. One of our main contributions is the derivation of novel effective equations for suitably renormalized unknowns. Interestingly, this renormalization is not performed with respect to a fixed background, but does involve the energy of the coupled system of wave equations. In addition, we construct an open set of initial data that are arbitrarily close to the expected asymptotic behavior. We emphasize that, in comparison, the class of Gowdy spacetimes exhibits a very different dynamical behavior to the one we uncover in the present work for general polarized T^2-symmetric spacetimes. Furthermore, all the conclusions of this paper are valid within the framework of weakly T^2-symmetric spacetimes previously introduced by the authors.

1. Introduction

This is the third of a series of papers [LeFloch and Smulevici 2015; 2016] devoted to the study of weakly regular, T^2-symmetric, vacuum spacetimes. There has been extensive work on the mathematical analysis of T^2-symmetric spacetimes with high regularity and we refer for instance to the introduction of [Smulevici 2011] for related literature. Our motivation in studying these spacetimes is two-fold. First of all, given the high degree of symmetry, one can study these solutions under much weaker regularity than in the general case. In [LeFloch and Smulevici 2015], we introduced the notion of weakly regular, T^2-symmetric, vacuum spacetime and we established a future-expanding, global existence theory in the so-called areal coordinates — generalizing a previous result in the smooth setup [Berger et al. 1997]. Our notion of weakly regular spacetimes extended a notion first proposed by Christodoulou [1993] (see also [LeFloch and Sormani 2015]) for radially symmetric spacetimes and later by [LeFloch 2015; LeFloch and Mardare 2007; LeFloch and Rendall 2011; LeFloch and Stewart 2005; LeFloch and Stewart 2011]

MSC2010: 83C05, 83C20, 35Q76.
Keywords: Einstein equations, T^2-symmetry, future expanding spacetime, late-time asymptotics, geodesic completeness.

for Gowdy symmetric spacetimes. See also the more recent developments in [Grubic and LeFloch 2013; 2015].

Our second motivation comes from the fact that, apart from special cases (see, for instance, [Chruściel et al. 1990; Ringström 2004; 2009]), a complete description of the late-time asymptotics of T^2-symmetric spacetimes has not been given yet *even for smooth initial data sets*. In fact, the techniques available until now provide the existence of future developments, but are not sufficient to prove that these spacetimes are future geodesically complete or not.

Recall that a T^2-symmetric, vacuum spacetime is a solution to the vacuum Einstein equations $\mathrm{Ric}(g) = 0$ arising from an initial data set which is assumed to be invariant under an action of the Lie group T^2. We are concerned here with the study of T^2-symmetric spacetime arising from initial data given on T^3. For such spacetimes, it is known [Chruściel 1990] that, unless the spacetime is flat (and therefore the solution is trivial) the area of the orbits of symmetry, say R, admits a timelike gradient and, therefore, can be used as time coordinate and leads one to define the so-called areal gauge. By convention, we can choose the time direction so that R increases toward the future. In the present paper, we restrict attention to *polarized* T^2-symmetric spacetimes, which are T^2-symmetric spacetimes for which the Killing fields generating the T^2 symmetry can be chosen to be mutually orthogonal.

Our main result is a complete description of the future time-asymptotics of polarized, T^2-symmetric, vacuum spacetimes, under the assumption that one starts sufficiently close to the expected asymptotic regime. As a consequence, it follows that these spacetimes are future geodesically complete. We refer to Theorems 7.1 and 8.1 for precise statements. These results are new even for smooth initial data, but we also emphasize that all of our estimates are valid within the framework of weakly regular, T^2-symmetric spacetimes introduced in [LeFloch and Smulevici 2015].

Prior to the present work, two important subclasses of T^2-symmetric solutions were studied in the literature. First of all, when the initial data set is invariant not only by an action of T^2 on T^3 but by the action of T^3 on itself, then the spacetime is homogeneous, i.e., admits three independent spatial Killing fields. The Einstein equations then reduce to a set of ordinary differential equations. Second, another subclass of solutions is the class of Gowdy spacetimes, which, by definition, are T^2-symmetric solutions for which the family of 2-planes orthogonal to the orbits of symmetry is integrable. One of the main differences between the Gowdy solutions and the general T^2-symmetric solutions is that the equations in areal gauge are semilinear in the Gowdy case, while they are quasilinear in general. The future time-asymptotics of Gowdy spacetimes were derived by Ringström [2004] (see also [Chruściel et al. 1990] for polarized Gowdy spacetimes).

The following question thus arises. Are the asymptotics of homogeneous T^2-symmetric or Gowdy spacetimes *stable within* the whole set of T^2-symmetric solutions? For homogeneous solutions, it turns out that there are not even stable within the class of Gowdy spacetimes [Ringström 2004]. As far as Gowdy spacetimes are concerned, the asymptotics derived in the present work show that they are not stable within the set of T^2-symmetric solutions. For instance, according to Theorem 7.1, the norm of the gradient of R behaves like R^{-2}, while it decays exponentially in the Gowdy case. Of course, one question which remains open is whether the future asymptotic behavior that we uncover here is stable, first within the whole class of T^2-symmetric solutions (i.e., for nonpolarized solutions) and, then, within the class of solutions arising from arbitrary initial data defined on T^3. We observe that many of the estimates we prove below can be generalized to the nonpolarized case.

Independently of this work, Ringström [2015] has recently obtained interesting and complementary results on T^2-symmetric spacetimes. His main results can be summarized as follows. For any T^2-symmetric spacetime that is nonflat and non-Gowdy, there is a certain geometric quantity[1] which, if bounded as $R \to +\infty$, implies that the solution is homogeneous. This result does not give sharp asymptotics on the solutions, but it is a large-data result and therefore, it is so far the strongest result available for T^2-symmetric spacetimes with arbitrary data. It implies, in particular, that the asymptotics of non-Gowdy, nonhomogeneous solutions are quite different from the asymptotics of homogeneous or Gowdy solutions. A second set of results proved in [Ringström 2015] concerns polarized T^2-symmetric under a smallness assumption (which is slightly different from the initial data assumption that we make here). A partial set of asymptotics is then obtained therein, while, in the present work, we derive a full set of late-time asymptotics; it is interesting to point out that the methods of proof appear to be quite different.

The rest of this paper is organized as follows. In the following section, we introduce standard material on T^2-symmetric and polarized solutions, which we will use throughout. In particular, we recall the global existence of areal foliation for weakly regular initial data established in [LeFloch and Smulevici 2015]. Apart from this result, this paper is essentially self-contained. We conclude the preliminary section by presenting the general strategy that we will use in order to derive the asymptotics. In Section 3, we derive some formulas for the evolution of certain mean values and we also provide some estimates about the commutator associated with the time derivative operator and the spatial average operator. Section 4 is devoted to the analysis of the corrected energy. In Section 5, we introduce several renormalized unknowns, derive a system of evolution equations for them and provide estimates on various error terms arising in the analysis. In Section 6, we introduce and close a small bootstrap argument, linking all the previous estimates together. In Sections 7 and 8, we present and give the proofs of the main results of this paper, concerning the full set of asymptotics and the geodesic completeness of these spacetimes, respectively. Finally, in Section 9, we construct an open set of initial data satisfying the assumptions of Theorem 7.1.

2. Preliminaries on T^2-symmetric polarized solutions

2A. ***Einstein equations in areal coordinates.*** Let $(\mathcal{M}, g)$ be a weakly regular, T^2-symmetric spacetime, understood in the sense introduced in [LeFloch and Smulevici 2015]. From the existence theory therein, we know that if $R : \mathcal{M} \to \mathbb{R}$ denotes the area of the orbits of the symmetry group then its gradient vector field ∇R is timelike (and future oriented thanks to the standard normalization adopted in [LeFloch and Smulevici 2015]) and, consequently, the area can be used as a time coordinate. In these *areal coordinates*, the variable R exhausts the interval $[R_0, +\infty)$, where $R_0 > 0$ is the (assumed) constant value of the area on the initial slice and the metric takes the form

$$g = e^{2(\eta-U)}(-dR^2 + a^{-2}\, d\theta^2) + e^{2U}(dx + A\, dy + (G + AH)\, d\theta)^2 + e^{-2U} R^2 (dy + H\, d\theta)^2. \tag{2-1}$$

[1] In the notation of this paper, it coincides with the quantity $\mathscr{P}$ introduced in (2-19).

Here, the independent variables x, y and θ belong to S^1 (the 1-dimensional torus or circle) and the metric coefficients U, A, η, a, G and H are functions of (R, θ), only. We will, for convenience in the presentation, identify S^1 with the interval $[0, 2\pi]$ and functions defined on S^1 with 2π-periodic functions. The vector fields ∂_x and ∂_y are Killing fields for the above metric and so are any linear combinations of ∂_x and ∂_y.

We are interested here in *polarized* T^2-symmetric spacetimes, defined as follows.

Definition 2.1. A T^2-symmetric spacetime is said to be polarized if one can choose linear combinations X and Y of the vector fields ∂_x and ∂_y generating the T^2 symmetry such that $g(X, Y) = 0$.

For a polarized spacetime, it follows that the metric can be rewritten (possibly after a change of the coordinates x and y) as

$$g = e^{2(\eta-U)}(-dR^2 + a^{-2}\,d\theta^2) + e^{2U}(dx + G\,d\theta)^2 + e^{-2U}R^2(dy + H\,d\theta)^2. \tag{2-2}$$

Now, the Einstein equations for T^2-symmetric spacetimes written in areal coordinates have been derived in [Berger et al. 1997] for smooth solutions (see also [Chruściel 1990] for the existence of areal time). In [LeFloch and Smulevici 2015], we introduced the *weak version of the Einstein equations* for weakly regular, T^2-symmetric spacetimes and we proved that, using areal coordinates, we could still reduce the Einstein equations to those obtained in [Berger et al. 1997]. In the polarized case, we are thus left with the following system of partial differential equations:

(1) Three evolution equations for the metric coefficients U, η and a:

$$(Ra^{-1}U_R)_R - (RaU_\theta)_\theta = 0, \tag{2-3}$$

$$(a^{-1}\eta_R)_R - (a\eta_\theta)_\theta = \Omega^\eta - \frac{1}{R^{3/2}}(R^{3/2}(a^{-1})_R)_R, \tag{2-4}$$

$$(2\ln a)_R = -\frac{K^2}{R^3}e^{2\eta}, \tag{2-5}$$

where K is a real constant and $\Omega^\eta := -a^{-1}U_R^2 + aU_\theta^2$.

(2) Two constraint equations for the metric coefficient η:

$$\eta_R + \frac{K^2}{4R^3}e^{2\eta} = aRE, \tag{2-6}$$

$$\eta_\theta = RF, \tag{2-7}$$

where $E := a^{-1}U_R^2 + aU_\theta^2$ and $F := 2U_RU_\theta$.

(3) Two equations for the twists:

$$G_R = 0 \quad \text{and} \quad H_R = \frac{K}{R^3}a^{-1}e^{2\eta}. \tag{2-8}$$

Here, K is the *twist constant* and $K = 0$ corresponds geometrically to the integrability of the family of 2-planes orthogonal to ∂_x and ∂_y. The special solutions with $K = 0$ are called Gowdy spacetimes (with

T^3 topology). Since the dynamics of Gowdy spacetimes are well known [Ringström 2004], we focus here exclusively on the case $K \neq 0$.

Note that the metric functions G and H do not appear in the equations apart from (2-8). These latter equations can simply be integrated in R, once enough information on their right-hand sides is obtained. They will therefore be ignored in most parts of this paper. Note also that (2-4) is actually a redundant equation, i.e., can be deduced from the other equations.[2]

Finally, observe that the identity

$$\left(\frac{e^{2\eta}}{a}\right)_R = 2REe^{2\eta} \tag{2-9}$$

will be useful later in this paper; it can be easily derived from the Einstein equations (2-5) and (2-6).

2B. *Global existence in areal coordinates.* In [LeFloch and Smulevici 2015], we proved local and global existence results for general T^2-symmetric spacetimes in areal coordinates. In the specific case of polarized, T^2-symmetric spacetimes, these results imply the following conclusion:

Theorem 2.2 (global existence theory in areal coordinates). *Fix any constants K, $R_0 > 0$. Consider any initial data $(U_0, U_1) \in H^1(S^1) \times L^2(S^1)$, $a_0 \in W^{2,1}(S^1)$ and $\eta_0 \in W^{1,1}(S^1)$ such that $a_0 > 0$. Suppose moreover that the constraint equation* (2-7) *is satisfied initially, i.e.,*

$$\partial_\theta(\eta_0) = 2R_0 U_1 \, \partial_\theta(U_0). \tag{2-10}$$

Let $\mathscr{C}$ be the class of functions (U, η, a) such that

$$\begin{aligned} &U \in C^1([R_0, +\infty), L^2(S^1)) \cap C^0([R_0, +\infty); H^1(S^1)), \\ &\eta \in C^0([R_0, +\infty); W^{1,1}(S^1)), \\ &a \in C^0([R_0, +\infty); W^{2,1}(S^1)). \end{aligned}$$

Then there exists a unique solution $(U, \eta, a) \in \mathscr{C}$ of the Einstein equations (2-3)–(2-7) *which assumes the given initial data at $R = R_0$, in the sense*

$$U(R_0) = U_0, \quad U_R(R_0) = U_1, \quad \eta(R_0) = \eta_0, \quad a(R_0) = a_0.$$

Moreover, on any compact time interval, the solution can be uniformly approximated by smooth solutions in the norm associated with $\mathscr{C}$.

Since all of our estimates here will be compatible with the density property stated at the end of the above theorem, it is sufficient to perform our analysis by assuming our initial data to be smooth.

---

[2]More precisely, (2-4) can be obtained by multiplying (2-6) and (2-7) by a^{-1} and a, respectively, differentiating the resulting equations in R and θ and taking their differences before replacing second derivatives of U and first derivatives of a using the evolution equations.

2C. ***Energy functionals.*** Important control on the metric coefficients, mostly on their first-order derivatives, is obtained by analyzing the energy functionals

$$\mathscr{E}(R) := \int_{S^1} E(R,\theta)\, d\theta, \quad E = a^{-1}U_R^2 + a\, U_\theta^2, \tag{2-11}$$

and

$$\mathscr{E}_K(R) := \int_{S^1} E_K(R,\theta)\, d\theta, \quad E_K := E + \frac{K^2}{4R^4} a^{-1} e^{2\eta}. \tag{2-12}$$

Using the Einstein equations (2-3)–(2-7), it follows that both functionals are *nonincreasing* in time, with

$$\begin{aligned} \frac{d}{dR}\mathscr{E}(R) &= -\frac{K^2}{2R^3}\int_{S^1} Ee^{2\eta}\, d\theta - \frac{2}{R}\int_{S^1} a^{-1}(U_R)^2\, d\theta, \\ \frac{d}{dR}\mathscr{E}_K(R) &= -\frac{K^2}{R^5}\int_{S^1} a^{-1}e^{2\eta}\, d\theta - \frac{2}{R}\int_{S^1} a^{-1}(U_R)^2\, d\theta. \end{aligned} \tag{2-13}$$

As a direct consequence, we have the following result:

Lemma 2.3 (uniform energy bounds for T^2-symmetric spacetimes). *The following uniform bounds hold*:

$$\sup_{R\in[R_0,+\infty)} \mathscr{E}(R) \le \mathscr{E}(R_0) \quad \textit{and} \quad \sup_{R\in[R_0,+\infty)} \mathscr{E}_K(R) \le \mathscr{E}_K(R_0), \tag{2-14}$$

as well as the spacetime bounds

$$\int_{R_0}^{+\infty}\int_{S^1} (a^{-1}c_0^U(U_R)^2 + ac_1^U(U_\theta)^2)\, dR\, d\theta \le \mathscr{E}(R_0), \tag{2-15}$$

$$\int_{R_0}^{+\infty} \frac{K^2}{R^5}\int_{S^1} e^{2\eta}\, a^{-1}\, dR\, d\theta \le \mathscr{E}_K(R_0), \tag{2-16}$$

with

$$c_0^U := \frac{2}{R} + \frac{K^2}{2R^3}e^{2\eta} \quad \textit{and} \quad c_1^U := \frac{K^2}{2R^3}e^{2\eta}.$$

2D. ***Heuristics and general strategy.*** To understand the asymptotic behavior of the solutions to wave equations such as (2-3), it is important to note that, while for the flat wave operator in $1+1$ dimensions there is no decay of solutions, the R-weights present in (2-3) reflect some expansion of our spacetime and that, in general, waves decay on expanding spacetimes.

The general strategy to capture this decay is to first observe that the global energy dissipation bound (2-15) associated with the energy functional $\mathscr{E}(R)$ gives an integrated energy decay estimate but with weaker weights for U_θ than for U_R (see the missing $2/R$ in c_1^U compared to c_0^U). To match the weights between U_R and U_θ, we will work instead with the *modified energy functional*

$$\widehat{\mathscr{E}}(R) := \mathscr{E}(R) + \mathscr{G}^U(R) \tag{2-17}$$

with

$$\mathscr{G}^U := \frac{1}{R}\int_{S^1} (U - \langle U\rangle) U_R a^{-1}\, d\theta,$$

in which the average $\langle f \rangle$ of a function $f = f(\theta)$ is not defined with respect to the flat measure $d\theta$ but with respect to a weighted measure $a^{-1}\, d\theta$, i.e.,

$$\langle f \rangle := \frac{\int_{S^1} f a^{-1}\, d\theta}{\int_{S^1} a^{-1}\, d\theta}. \tag{2-18}$$

Our strategy is then to "trade" a time derivative for a space derivative. This method of proof was previously used in [Ringström 2004; Choquet-Bruhat and Moncrief 2001; Choquet-Bruhat 2003].

The following notation will be useful. We introduce the length $\mathscr{P}$ of the circle S^1 with respect to the measure $a^{-1}\, d\theta$, that is,

$$\mathscr{P}(R) := \int_0^{2\pi} a^{-1}\, d\theta, \tag{2-19}$$

which we refer to as the *perimeter*. The geometric interpretation of this quantity is that the principal symbol of the wave operator appearing in the wave equation (2-3) for U is that of the 2-dimensional metric

$$ds^2 = -dR^2 + a^{-2}\, d\theta^2.$$

Thus, $\mathscr{P}$ is the volume of the constant-R slice for this metric.

Naively, one may expect the following behavior as $R \to +\infty$. In view of the energy identity (2-13) satisfied by $\mathscr{E}$ and focusing on the second integral term, one may expect that

$$\frac{d}{dR}\mathscr{E} \le -\frac{2}{R}\mathscr{E} \quad \text{(modulo higher-order terms)},$$

so that $\mathscr{E}$ should decay like $1/R^2$. This behavior is indeed correct for spatially homogeneous spacetimes, as can be checked directly. However, for nonspatially homogeneous solutions, a space derivative must be recovered from a time derivative, using the corrected energy $\widehat{\mathscr{E}}$ defined in (2-17), as we already explained above. This would lead to a rate of decay determined by

$$\frac{d}{dR}\widehat{\mathscr{E}}(R) \le -\frac{1}{R}\widehat{\mathscr{E}}(R) \quad \text{(modulo higher-order terms)},$$

so that $\widehat{\mathscr{E}}$ should decay like $1/R$. If one can then check that the correction term in $\widehat{\mathscr{E}}$ is of order $o(1/R)$, it should follow that $\mathscr{E}(R)$ is of order $1/R$. This is indeed the rate of decay established by Ringström [2004] for (sufficiently regular) Gowdy spacetimes.

For the more general class of spacetimes under consideration in the present paper, and due to the variation of the metric coefficients a and η, the behavior $\mathscr{E} \sim 1/R$ is *not* consistent with the field equations, as we now check formally. At this stage of the discussion, we are working under the (later invalidated, below) assumption that the first term in (2-13) is negligible, say specifically

$$\frac{\|e^{2\eta}\|_{L^\infty(S_1)}}{2R^3} \lesssim \frac{1}{R^{1+\epsilon}}, \quad \epsilon > 0. \tag{2-20}$$

From (2-5) we would deduce

$$(\ln a)_R = -K^2 \frac{e^{2\eta}}{2R^3} \in L^1_R,$$

hence the coefficient a would then admit a finite limit as $R \to +\infty$. Next, in view of

$$\eta_R = -\frac{K^2}{2}\frac{e^{2\eta}}{2R^3} + aRE,$$

in which $\int_{S^1} RE\, d\theta$ is bounded thanks to our energy assumption, it would then follow that $\int_{S^1} \eta_R$ behaves like 1 and thus $\int_{S_1} \eta \sim R$ (modulo a multiplicative constant). In turn, this invalidates our original assumption (2-20).

This means that the first term in (2-13) should not be neglected and that it contributes significantly to the energy decay. We will prove that, modulo an error term due to the spatial variation of η, this term can be rewritten as $-(\mathscr{P}_R/\mathscr{P})\mathscr{E}$, where $\mathscr{P}$ is the perimeter defined by (2-19).

Taking this into account, it follows, assuming that all the error terms can be controlled, that the *rescaled energy*

$$\mathscr{F} := \mathscr{P}\widehat{\mathscr{E}} \tag{2-21}$$

should decay like $1/R$ and, in other words, the energy $\widehat{\mathscr{E}}$ should decay like $1/\mathscr{P}R$. This brings more decay into our analysis, provided the perimeter $\mathscr{P}$ is *growing as* $R \to +\infty$ — as we will actually show later. Indeed, we will establish that the perimeter and metric coefficients have the asymptotic behavior (possibly up to multiplicative constants)

$$\mathscr{P}(R) \sim R^{1/2}, \quad \mathscr{P}_R(R) \sim R^{-1/2}, \quad e^{2\eta} \sim R^2, \quad a \sim R^{-1/2}. \tag{2-22}$$

For the energy, we will therefore have $\mathscr{E} \sim R^{-3/2}$. Surprisingly, all the multiplicative constants in the above asymptotic behavior are linked to each other. For instance, we will show that $R^2\mathscr{P}^{-1}\mathscr{E} \to \frac{5}{4}$ as $R \to +\infty$. One of the main difficulties lies in fact in trying to understand these relations. Thus, our work really consists of three ingredients:

(1) A version of the corrected energy functionals adapted to polarized, T^2-symmetric spacetimes (Sections 3 and 4).

(2) A derivation and analysis of a dynamical system to understand the interplay between $\mathscr{P}$ and the energy functionals (Section 5).

(3) Estimates on all the error terms involved in the above two steps and the interplay between all the previous estimates. Since all the estimates involved in the above estimates depend on each other, we use a small bootstrap argument to obtain closure (Section 6).

Once these elements have been obtained, deriving the asymptotics of the solutions consists mostly in revisiting the previous estimates in the proper order (see Section 7). Finally, we prove the geodesic completeness by using the approach already developed in [LeFloch and Smulevici 2016] (see Section 8).

3. Evolution of the mean values

3A. ***The length variable.*** In addition to the perimeter $\mathscr{P}(R)$ introduced in (2-19), the metric coefficient a also determines a *length function*

$$\vartheta(\theta, R) := \int_0^\theta a^{-1}\, d\theta, \quad \theta \in S^1, \tag{3-1}$$

and its inverse $\Theta = \Theta(\vartheta, R)$ (for each fixed R). In other words, we set $\Theta(\vartheta(\theta, R), R) = \theta$ for all $\theta \in S^1$, so that

$$\Theta(\vartheta, R) = \int_0^\vartheta a(\Theta(\vartheta', R), R)\, d\vartheta', \quad \Theta(\mathscr{P}(R), R) = 2\pi. \tag{3-2}$$

Using the change of variable determined by the length function, we can parameterize any function $f = f(R, \theta)$ into $\tilde{f} = \tilde{f}(R, \vartheta)$, defined by

$$\tilde{f}(R, \vartheta) := f(R, \Theta(\vartheta, R)). \tag{3-3}$$

This is nothing but a change of coordinates from (R, θ) to (R, ϑ), but we insist on keeping the "tilde notation" in order to avoid confusion (when taking averages and R-derivatives).

The average of any $L^1(S^1)$ function f is now naturally computed with respect to the measure $d\vartheta$, that is,

$$\langle \tilde{f}(R)\rangle := \frac{1}{\mathscr{P}(R)} \int_0^{\mathscr{P}(R)} \tilde{f}(R)\, d\vartheta. = \frac{1}{\mathscr{P}(R)} \int_0^{2\pi} f(R) a(R)^{-1}\, d\theta = \langle f(R)\rangle, \tag{3-4}$$

which, as stated, obviously coincides with $\langle f(R)\rangle$ as defined by (2-18). Note that the periodicity property is preserved in the new variable, that is,

$$\tilde{f}(R, \vartheta + \mathscr{P}(R)) = \tilde{f}(R, \vartheta)$$

for all relevant values of R and ϑ.

Using the above notation, we can for instance rewrite the correction $\mathscr{G}^U$ introduced in (2-17) in the form

$$\mathscr{G}^U(R) := \frac{1}{R} \int_0^{\mathscr{P}(R)} \big(\widetilde{U}(R) - \langle \widetilde{U}(R)\rangle\big) \widetilde{U}_R(R)\, d\vartheta. \tag{3-5}$$

This form has some advantages when differentiating with respect to R, since it directly involves the perimeter and its derivative, which have a geometric meaning.

3B. ***Derivatives of the mean values.*** We will be taking time derivatives of the above quantities but, since the time-derivative operator and the spatial averaging operator do not commute, an analysis of the corresponding "commutator" will be required. The following properties will be used throughout the rest of this article.

Lemma 3.1 (general identities for the mean values). *For any* (*sufficiently regular*) *function* $f = f(R, \theta)$, *one has*

$$\frac{d}{dR}\langle\tilde{f}\rangle = \langle\tilde{f}_R\rangle + \frac{K^2}{2R^3}\langle\widetilde{fe^{2\eta}}\rangle - \frac{\mathscr{P}_R}{\mathscr{P}}\langle\tilde{f}\rangle,$$

$$\frac{d}{dR}(\mathscr{P}\langle\tilde{f}\rangle) = \mathscr{P}\langle\tilde{f}_R\rangle + \mathscr{P}\frac{K^2}{2R^3}\langle\widetilde{fe^{2\eta}}\rangle,$$

in which $\tilde{f}$ *is defined by* (3-3).

Proof. From the definition

$$\langle\tilde{f}\rangle = \frac{1}{\mathscr{P}}\int_0^{\mathscr{P}} \tilde{f}\, d\vartheta = \frac{1}{\mathscr{P}}\int_0^{2\pi} f(R,\theta)a^{-1}\, d\theta,$$

we deduce that

$$\begin{aligned}\frac{d}{dR}\langle\tilde{f}\rangle &= \langle\tilde{f}_R\rangle + \frac{1}{\mathscr{P}}\int_0^{2\pi} f(a^{-1})_R\, d\theta - \frac{\mathscr{P}_R}{\mathscr{P}}\langle\tilde{f}\rangle \\ &= \langle\tilde{f}_R\rangle + \frac{1}{\mathscr{P}}\int_0^{2\pi} f\frac{K^2e^{2\eta}}{2R^3}a^{-1}\, d\theta - \frac{\mathscr{P}_R}{\mathscr{P}}\langle\tilde{f}\rangle \\ &= \langle\tilde{f}_R\rangle + \frac{K^2}{2R^3}\langle\widetilde{fe^{2\eta}}\rangle - \frac{\mathscr{P}_R}{\mathscr{P}}\langle\tilde{f}\rangle,\end{aligned}$$

which leads us to the two identities stated in the lemma. □

The above lemma allows us to derive the following estimate:

Lemma 3.2 (commutator estimate). *The commutator associated with the time-differentiation and averaging operators satisfies*, *for all functions* f,

$$\left|\frac{d}{dR}\langle\tilde{f}\rangle - \langle\tilde{f}_R\rangle\right| \le \frac{\pi K^2}{R^3}\langle|\tilde{f}|\rangle\|(e^{2\eta})_\theta\|_{L^1(S^1)}.$$

Proof. From the above lemma, the expression of $\mathscr{P}_R$ and the evolution equation satisfied by a, we deduce

$$\begin{aligned}\left|\frac{d}{dR}\langle\tilde{f}\rangle - \langle\tilde{f}_R\rangle\right| &\le \frac{K^2}{2R^3\mathscr{P}^2}\int_0^{2\pi}|f|a^{-1}(R,\theta)\left|e^{2\eta}(R,\theta)\mathscr{P} - \int_0^{2\pi} e^{2\eta}a^{-1}(R,\theta')\, d\theta'\right| d\theta \\ &\le \frac{\pi K^2}{R^3\mathscr{P}}\langle|\tilde{f}|\rangle \sup_{\theta\in S^1}\left|e^{2\eta}(R,\theta)\mathscr{P} - \int_0^{2\pi} e^{2\eta}(R,\theta')a^{-1}(R,\theta')\, d\theta'\right|\end{aligned}$$

with

$$\sup_{\theta\in S^1}\left|e^{2\eta}(R,\theta)\mathscr{P} - \int_0^{2\pi} e^{2\eta}(R,\theta')a^{-1}(R,\theta')\, d\theta'\right| \le \mathscr{P}\big(\sup_{S^1} e^{2\eta} - \min_{S^1} e^{2\eta}\big) \le \mathscr{P}\|(e^{2\eta})_\theta\|_{L^1(S^1)}. \quad \square$$

The following conserved quantity will also be useful in our analysis. It follows simply after a global integration in space of the wave equation (2-3) and an integration in R on $[R_1, R]$.

Lemma 3.3. *For all* $R \ge R_1$, *the following conservation law holds*:

$$R\mathscr{P}\langle\widetilde{U}_R\rangle = R_1\mathscr{P}(R_1)\langle\widetilde{U}_R\rangle(R_1).$$

4. Evolution of the modified energy functional

4A. ***Evolution of correction terms.*** Using Lemma 3.1, we can compute the time derivative of the corrector $\mathscr{G}^U$ in (3-5); indeed:

$$\begin{aligned}\frac{d}{dR}\mathscr{G}^U &= -\frac{1}{R}\mathscr{G}^U + \frac{1}{R}\left(\int_0^{2\pi} (U - \langle\widetilde{U}\rangle)U_R a^{-1}\, d\theta\right)_R \\ &= -\frac{1}{R}\mathscr{G}^U + \frac{1}{R}\int_0^{2\pi} U_R^2 a^{-1}\, d\theta + \frac{1}{R}\int_0^{\mathscr{P}} \left(-\langle\widetilde{U}_R\rangle - \frac{K^2}{2R^3}\langle\widetilde{Ue^{2\eta}}\rangle + \frac{\mathscr{P}_R}{\mathscr{P}}\langle\widetilde{U}\rangle\right)\widetilde{U}_R\, d\vartheta \\ &\qquad + \frac{1}{R}\int_0^{2\pi} (U - \langle\widetilde{U}\rangle)(U_R a^{-1})_R\, d\theta,\end{aligned}$$

so that, by using the field equation (2-3) satisfied by U,

$$\begin{aligned}\frac{d}{dR}\mathscr{G}^U &= -\frac{1}{R}\mathscr{G}^U + \frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_R^2\, d\vartheta - \frac{\mathscr{P}}{R}(\langle\widetilde{U}_R\rangle)^2 - \frac{K^2}{2R^4}\mathscr{P}\langle\widetilde{U}_R\rangle\langle\widetilde{Ue^{2\eta}}\rangle \\ &\qquad + \frac{\mathscr{P}_R}{R}\langle\widetilde{U}\rangle\langle\widetilde{U}_R\rangle + \frac{1}{R}\int_0^{2\pi} (U - \langle\widetilde{U}\rangle)\left(-\frac{U_R a^{-1}}{R} + (aU_\theta)_\theta\right) d\theta.\end{aligned}$$

Integrating by parts the last term, we obtain

$$-\frac{2}{R}\mathscr{G}^U + \frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_R^2\, d\vartheta - \frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_\vartheta^2\, d\vartheta - \frac{\mathscr{P}}{R}(\langle\widetilde{U}_R\rangle)^2 - \frac{K^2}{2R^4}\mathscr{P}\langle\widetilde{U}_R\rangle\langle\widetilde{Ue^{2\eta}}\rangle + \frac{\mathscr{P}_R}{R}\langle\widetilde{U}\rangle\langle\widetilde{U}_R\rangle.$$

After reorganizing some of the terms, this leads us to

$$\frac{d}{dR}\mathscr{G}^U = -\frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_\vartheta^2\, d\vartheta + \frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_R^2\, d\vartheta - \frac{1}{R}\mathscr{G}^U - \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{G}^U + \Omega_{\mathscr{G}^U}, \tag{4-1}$$

with

$$\Omega_{\mathscr{G}^U} = \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{G}^U - \frac{\mathscr{P}}{R}(\langle\widetilde{U}_R\rangle)^2 - \frac{1}{R}\mathscr{G}^U - \frac{K^2}{2R^4}\mathscr{P}\langle\widetilde{U}_R\rangle\langle\widetilde{Ue^{2\eta}}\rangle + \frac{\mathscr{P}_R}{R}\langle\widetilde{U}\rangle\langle\widetilde{U}_R\rangle. \tag{4-2}$$

The term $\Omega_{\mathscr{G}^U}$ will be shown to be an "error term", while the remaining terms in the right-hand side of (4-1) will contribute to the derivation of a sharp energy decay estimate. In (4-1) and (4-2), we have added and subtracted the term $(\mathscr{P}_R/\mathscr{P})\mathscr{G}^U$, as this will simplify some of our estimates.

4B. ***Evolution of the corrected energy.*** Summing together the contributions of the energy and the correction $\mathscr{G}^U$, we find

$$\begin{aligned}\frac{d}{dR}(\mathscr{E} + \mathscr{G}^U) &= -\frac{K^2}{2R^3}\int_0^{2\pi} Ee^{2\eta}\, d\theta - \frac{2}{R}\int_{S^1} (a^{-1}U_R^2)\, d\theta \\ &\qquad + \frac{1}{R}\int_0^{2\pi} a^{-1}U_R^2\, d\theta - \frac{1}{R}\int_0^{\mathscr{P}} \widetilde{U}_\vartheta^2\, d\vartheta - \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{G}^U - \frac{1}{R}\mathscr{G}^U + \Omega_{\mathscr{G}^U} \\ &= -\frac{\mathscr{P}_R}{\mathscr{P}}(\mathscr{E} + \mathscr{G}^U) - \frac{1}{R}(\mathscr{E} + \mathscr{G}^U) + \Omega_{\mathscr{E}} + \Omega_{\mathscr{G}^U},\end{aligned} \tag{4-3}$$

where the error terms are $\Omega_{\mathscr{G}^U}$, defined by (4-2), and

$$\Omega_{\mathscr{E}} = \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{E} - \frac{K^2}{2R^3}\int_0^{2\pi} Ee^{2\eta}\,d\theta.$$

4C. ***Estimate for the energy correction.*** We will need the following 1-dimensional Poincaré (or Wirtinger) inequality: for any $a > 0$, if f is an a-periodic function in $H^1(0, a)$ and has mean value 0 on this interval, then

$$\int_{[0,a]} f^2 \le \frac{a^2}{4\pi^2}\int_{[0,a]} f'^2. \tag{4-4}$$

This is easily checked by, for instance, using a Fourier decomposition of f. Using the above notation, we have the following lemma:

Lemma 4.1 (estimate of the $\mathscr{G}^U$ correction of the energy). *We have*

$$|\mathscr{G}^U(R)| \le \frac{\mathscr{P}(R)}{4\pi R}\mathscr{E}(R).$$

Proof. We apply the inequality $ab \le \frac{1}{2}(a^2+b^2)$ to the integrand of $R\mathscr{G}^U$, but we insert weights of $\mathscr{P}/(2\pi)$ so as to obtain

$$|R\mathscr{G}^U| \le \frac{\mathscr{P}}{4\pi}\int_0^{\mathscr{P}} \widetilde{U}_R^2\,d\vartheta + \frac{2\pi}{2\mathscr{P}}\int_0^{\mathscr{P}} (\widetilde{U} - \langle\widetilde{U}\rangle)^2\,d\vartheta \le \frac{\mathscr{P}}{4\pi}\int_0^{2\pi} U_R^2 a^{-1}\,d\theta + \frac{\mathscr{P}}{4\pi}\int_0^{\mathscr{P}} \widetilde{U}_\vartheta^2\,d\vartheta = \frac{\mathscr{P}}{4\pi}\mathscr{E}. \quad \square$$

4D. ***Estimates for the error terms.*** In this section, we estimate all the error arising in the corrected energy formula (4-3).

Lemma 4.2 (estimate for the $|\Omega_{\mathscr{E}}|$ error term). *We have*

$$|\Omega_{\mathscr{E}}| \le \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi} 2REe^{2\eta} = \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R.$$

Proof. Recall that

$$\frac{\mathscr{P}_R}{\mathscr{P}} = \frac{K^2}{2R^3}\int_0^{2\pi} e^{2\eta}a^{-1}\,d\theta\left(\int_0^{2\pi} a^{-1}\,d\theta\right)^{-1},$$

so that

$$\begin{aligned}\left|-\frac{K^2}{2R^3}\int_0^{2\pi} Ee^{2\eta}\,d\theta + \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{E}\right| &\le \frac{K^2}{2R^3\mathscr{P}}\int_0^{2\pi} E(R,\theta)\,d\theta\int_0^{2\pi} a^{-1}(R,\theta')|e^{2\eta}(R,\theta') - e^{2\eta}(R,\theta)|\,d\theta' \\ &\le \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi} |2\eta_\theta|e^{2\eta}\,d\theta \\ &\le \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi} 2REe^{2\eta}\,d\theta = \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R d\theta,\end{aligned}$$

where we have used the constraint equation (2-7) for η_θ and the identity (2-9). $\square$

Next, we analyze the error term $\Omega_{\mathscr{G}^U}$. It is convenient to split it into three components as follows: $\Omega_{\mathscr{G}^U} = I_1 + I_2 + I_3$, where I_1, I_2 and I_3 are defined as

$$
\begin{aligned}
I_1 &= -\frac{1}{R}\mathscr{G}^U, \\
I_2 &= \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{G}^U + \frac{\mathscr{P}_R}{R}\langle\widetilde{U}\rangle\langle\widetilde{U}_R\rangle - \frac{K^2}{2R^4}\mathscr{P}\langle\widetilde{U}_R\rangle\langle\widetilde{Ue^{2\eta}}\rangle, \\
I_3 &= -\frac{\mathscr{P}}{R}(\langle\widetilde{U}_R\rangle)^2.
\end{aligned}
$$

Lemma 4.3. *The following estimates hold*:

$$
\begin{aligned}
|I_1| &\le \frac{\mathscr{P}(R)}{4\pi R^2}\mathscr{E}(R), \\
|I_2| &\le \frac{\mathscr{P}_R}{R}\mathscr{E}, \\
|I_3| &\le \frac{\mathscr{A}}{R^3\mathscr{P}(R)},
\end{aligned}
$$

where $\mathscr{A}$ is a nonnegative constant determined by the initial data:

$$
\mathscr{A} = R_1^2\mathscr{P}(R_1)^2(\langle\widetilde{U}_R\rangle)^2(R_1).
$$

Proof. The estimates on I_1 and I_3 follow immediately from Lemmas 4.1 and 3.3, respectively. We then estimate I_2 as follows. Note first that

$$
\begin{aligned}
I_2 &= \frac{\mathscr{P}_R}{\mathscr{P}}\mathscr{G}^U + \frac{\mathscr{P}_R}{R}\langle\widetilde{U}\rangle\langle\widetilde{U}_R\rangle - \frac{K^2}{2R^4}\mathscr{P}\langle\widetilde{U}_R\rangle\langle\widetilde{Ue^{2\eta}}\rangle \\
&= \frac{K^2}{2R^4\mathscr{P}}\int_0^{2\pi} U_R(R,\theta')a^{-1}(R,\theta')\left(\int_0^{2\pi} e^{2\eta}(R,\theta)a^{-1}(R,\theta)[U(R,\theta') - U(R,\theta)]\,d\theta\right)d\theta';
\end{aligned}
$$

hence,

$$
|I_2| \le \frac{K^2}{2R^4\mathscr{P}}\int_0^{\mathscr{P}}|\widetilde{U}_R|\,d\vartheta\int_0^{\mathscr{P}} e^{2\tilde{\eta}}\,d\vartheta\int_0^{\mathscr{P}}|\widetilde{U}_\vartheta|\,d\vartheta \le \frac{\mathscr{P}_R}{R\mathscr{P}}(\mathscr{E}^{1/2}\mathscr{P}^{1/2})^2 \le \frac{\mathscr{P}_R}{R}\mathscr{E}. \qquad \square
$$

4E. ***Combining the estimates for the corrected energy.*** Collecting all the estimates for the error terms above and noting that I_3 has a sign, we obtain the estimate

$$
\frac{d}{dR}(\mathscr{E}+\mathscr{G}^U) + \left(\frac{1}{R}+\frac{\mathscr{P}_R}{\mathscr{P}}\right)(\mathscr{E}+\mathscr{G}^U) \le \frac{\mathscr{P}}{4\pi R^2}\mathscr{E} + \frac{\mathscr{P}_R}{R}\mathscr{E} + \mathscr{E}\frac{K^2}{2R^3}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R,
$$

from which it follows that

$$
\begin{aligned}
&R\mathscr{P}(\mathscr{E}+\mathscr{G}^U)(R) \\
&\quad \le R_0\mathscr{P}(\mathscr{E}+\mathscr{G}^U)(R_0) + \int_{R_0}^R \frac{\mathscr{P}^2\mathscr{E}}{4\pi R'}\,dR' + \int_{R_0}^R \mathscr{P}_R\mathscr{P}\mathscr{E}\,dR' + \int_{R_0}^R \mathscr{P}\mathscr{E}\frac{K^2}{2R'^2}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R d\theta\,dR'.
\end{aligned}
$$

Similarly, we can obtain

$$\frac{d}{dR}(R\mathscr{P}(\mathscr{E}+\mathscr{G}^U)) \geq -\frac{\mathscr{A}}{R^2} - \frac{\mathscr{P}^2}{4\pi R}\mathscr{E} + \mathscr{P}\mathscr{P}_R\mathscr{E} + \mathscr{P}\mathscr{E}\frac{K^2}{2R^2}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R, \tag{4-5}$$

leading to

$$R\mathscr{P}(\mathscr{E}+\mathscr{G}^U)(R) \geq R_0\mathscr{P}(\mathscr{E}+\mathscr{G}^U)(R_0) - \int_{R_0}^{R}\frac{\mathscr{P}^2\mathscr{E}}{4\pi R'}\,dR' - \int_{R_0}^{R}\mathscr{P}_R\mathscr{P}\mathscr{E}\,dR' - \int_{R_0}^{R}\mathscr{P}\mathscr{E}\frac{K^2}{2R'^2}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R d\theta\,dR' - \int_{R_0}^{R}\frac{\mathscr{A}}{R'^2}\,dR',$$

where $\mathscr{A}$ is the constant in Lemma 4.3.

5. A dynamical system for the renormalized unknowns

5A. ***The dynamical system.*** In the previous section, we have obtained differential inequalities for the quantity $\mathscr{P}(\mathscr{E}+\mathscr{G}^U)$, with error terms depending mostly on $\mathscr{E}$ and $\mathscr{P}$. In this section, we will try to obtain effective equations in order to control the asymptotic behavior of $\mathscr{P}$. For convenience, we introduce the notation

$$\mathcal{F} := \mathscr{P}\mathscr{E} \quad \text{and} \quad \mathcal{G} := \mathscr{P}(\mathscr{E}+\mathscr{G}^U).$$

We have thus seen that $\mathcal{G}$ satisfies "good" differential inequalities while it is ultimately $\mathcal{F}$ that we want to control, as it is a manifestly coercive quantity (contrary to $\mathcal{G}$). We will rely on the guess that the function $\mathcal{G}$ decays like $1/R$, but we will not use yet the differential inequalities derived for $\mathcal{G}$ in the previous section. In fact, $\mathcal{G}$ will appear here only in the form $R\mathcal{G}'/\mathcal{G}$.

The system of ODEs: spatial integration and first error terms. Let $\mathfrak{Q} = \int_{S^1}\frac{1}{2}K^2e^{2\eta}a^{-1}\,d\theta$. After integration in the spatial variable of the Einstein equations (2-5)–(2-6), we obtain

$$\mathscr{P}_R = \frac{\mathfrak{Q}}{R^3}, \tag{5-1}$$

$$\mathfrak{Q}_R = 2R\mathcal{F}\mathfrak{Q}\mathscr{P}^{-2} + \Omega_{\mathfrak{Q}}, \tag{5-2}$$

where $\Omega_{\mathfrak{Q}}$ is given by

$$\Omega_{\mathfrak{Q}} = 2R\left(\int_{S^1}\frac{K^2}{2}Ee^{2\eta} - \mathscr{P}^{-1}\mathscr{E}\mathfrak{Q}\right).$$

As in Lemma 4.2, $\Omega_{\mathfrak{Q}}$ satisfies the estimate

$$|\Omega_{\mathfrak{Q}}| \leq RK^2\mathscr{E}\int_0^{2\pi}\left(\frac{e^{2\eta}}{a}\right)_R d\theta = 2R\mathscr{E}\mathfrak{Q}_R. \tag{5-3}$$

Renormalization. According to our previous discussion, we expect $\mathscr{P}$ to blow-up in the limit. One can check heuristically that "$\mathscr{P}$ growing like $R^{1/2}$" and "$\mathfrak{Q}$ growing like $R^{5/2}$" seem the only possibilities (as powers of R) compatible with the equations, under the assumption that $\mathscr{P}\mathscr{E}$ behaves like R^{-1} (see the discussion at the end of Section 2D). Thus, one may try to introduce variables $\tilde{c} = \mathscr{P} R^{-1/2}$ and $\tilde{d} = \mathscr{Q} R^{-5/2}$ and prove that $\tilde{c}$ and $\tilde{d}$ converge to some finite values. Using (5-2), the equation for $\tilde{d}$ is then

$$\tilde{d}_R = \frac{\tilde{d}}{R}\Big(2R^2 \mathscr{F} \mathscr{P}^{-2} - \frac{5}{2}\Big) + \Omega_{\mathscr{Q}}.$$

From this equation and the coupled equation for $\tilde{c}$, it is not clear whether $\tilde{c}$ and $\tilde{d}$ converge. However, assuming $\Omega_{\mathscr{Q}}$ to be a negligible term, it suggests that $2R\mathscr{F}\mathscr{P}^{-2} \to \frac{5}{2}$ as $R \to +\infty$. Equivalently, it suggests that $\mathscr{P}/R\mathscr{F}^{1/2} \to \frac{2}{\sqrt{5}}$. Similarly, one can guess that $\mathscr{Q}/(R^3\mathscr{F}^{1/2}) \to \frac{1}{\sqrt{5}}$. We thus introduce a new set of variables c and d, replacing $\mathscr{P}$ and $\mathscr{Q}$, based on these considerations.

However, since it is actually $\mathscr{G}$ that satisfies "good" differential inequalities, we define c and d as

$$c := \frac{\mathscr{P}}{R\sqrt{\mathscr{G}}}, \tag{5-4}$$

$$d := \frac{\mathscr{Q}}{R^3\sqrt{\mathscr{G}}}, \tag{5-5}$$

where we recall that $\mathscr{G} = \mathscr{P}(\mathscr{E} + \mathscr{G}^U)$. Once again, we emphasize that, while $\mathscr{G}$ behaves asymptotically as $\mathscr{F}$, it is important to use this normalization rather than the one based on $\mathscr{F}$, since the normalization procedure will introduce a derivative of $\mathscr{G}$ in the equation and it is this derivative (rather than that of $\mathscr{F}$) that we can control directly.

Note that, while $\mathscr{F}$ is manifestly nonnegative, this is not the case for $\mathscr{G}$. In the rest of this section, we will assume that $\mathscr{G} > 0$, which ensures that all the computations below (as well as the definitions of c and d) make sense. In the next section, a lower bound on $\mathscr{G}$ using a bootstrap argument will be recovered.

An easy computation shows that (c,d) satisfies

$$c' = \frac{d}{R} - \frac{c}{R} - \frac{c}{2}\frac{\mathscr{G}'}{\mathscr{G}}, \tag{5-6}$$

$$d' = \frac{\mathscr{F}}{\mathscr{G}}\frac{2dc^{-2}}{R} - \frac{3}{R}d - \frac{d}{2}\frac{\mathscr{G}'}{\mathscr{G}} + \frac{\Omega_{\mathscr{Q}}}{R^3\sqrt{\mathscr{G}}}. \tag{5-7}$$

To find the correct limits for (c,d), let us first consider, the ordinary differential system

$$\begin{aligned} c' &= \frac{d}{R} - \frac{c}{R} + \frac{c}{2R}, \\ d' &= \frac{2dc^{-2}}{R} - \frac{3}{R}d + \frac{d}{2R}, \end{aligned} \tag{5-8}$$

which is obtained from the previous one by replacing $\mathscr{F}/\mathscr{G}$ by 1, dropping the error term $\Omega_{\mathscr{Q}}/(R^3\sqrt{\mathscr{G}})$ and replacing $-\mathscr{G}'/\mathscr{G}$ by $1/R$.

Looking now for a static point (c_∞, d_∞) of the above system, we find that there is only one solution: $c_\infty = \frac{2}{\sqrt{5}}$ and $d_\infty = \frac{1}{\sqrt{5}}$. Thus, let us introduce c_1 and d_1 by

$$c_1 = c - \frac{2}{\sqrt{5}} \quad \text{and} \quad d_1 = d - \frac{1}{\sqrt{5}}. \tag{5-9}$$

We finally deduce the equations satisfied by c_1 and d_1 from the equations (5-6)–(5-7); that is,

$$c_1' = \frac{d_1 + \frac{1}{\sqrt{5}}}{R} - \frac{\frac{2}{\sqrt{5}} + c_1}{R} - \frac{\frac{2}{\sqrt{5}} + c_1}{2}\frac{\mathcal{G}'}{\mathcal{G}}, \tag{5-10}$$

$$d_1' = \frac{\mathcal{F}}{\mathcal{G}}\frac{2}{R}\frac{d_1 + \frac{1}{\sqrt{5}}}{\left(\frac{2}{\sqrt{5}} + c_1\right)^2} - \frac{3}{R}\left(d_1 + \frac{1}{\sqrt{5}}\right) - \left(d_1 + \frac{1}{\sqrt{5}}\right)\frac{\mathcal{G}'}{2\mathcal{G}} + \frac{\Omega_{\mathcal{Q}}}{R^3\sqrt{\mathcal{G}}}. \tag{5-11}$$

Looking first at (5-10), we rewrite it in the form

$$c_1' = \frac{1}{R}d_1 - \frac{1}{2}\frac{c_1}{R} - \frac{c_1}{2R}\left(1 + R\frac{\mathcal{G}'}{\mathcal{G}}\right) - \frac{1}{R\sqrt{5}}\left(1 + R\frac{\mathcal{G}'}{\mathcal{G}}\right).$$

From (5-11), elementary calculations (keeping in mind the linearization of the system) lead us to

$$d_1' = -\frac{5}{2R}c_1 + \frac{d_1}{R}\left(-\frac{1}{2} - \frac{\mathcal{G}'}{2\mathcal{G}}R\right) - \frac{1}{R}\frac{1}{2\sqrt{5}}\left(1 + \frac{R\mathcal{G}'}{\mathcal{G}}\right) + \frac{1}{R\left(c_1 + \frac{2}{\sqrt{5}}\right)^2} f(d_1, c_1) + \frac{2}{R}\left(\frac{\mathcal{F}}{\mathcal{G}} - 1\right)\frac{d_1 + \frac{1}{\sqrt{5}}}{\left(\frac{2}{\sqrt{5}} + c_1\right)^2} + \frac{\Omega_{\mathcal{Q}}}{R^3\sqrt{\mathcal{G}}},$$

where $f(c_1, d_1)$ is a polynomial in c_1 and d_1 with vanishing linear part (the first terms are quadratic in c_1 and d_1). Thus, we have

$$d_1' = -\frac{5}{2R}c_1 + \Omega^d_{\text{lin}} + \Omega^d_1 + \Omega^d_2 + \Omega^d_3 + \Omega^d_4, \tag{5-12}$$

$$c_1' = \frac{d_1}{R} - \frac{c_1}{2R} + \Omega^c_{\text{lin}} + \Omega^c_1, \tag{5-13}$$

where the terms $\Omega_i^{c,d}$ contain all the error terms, i.e.,

$$\Omega^d_{\text{lin}} = -\frac{d_1}{2R}\left(1 + \frac{\mathcal{G}'}{\mathcal{G}}R\right), \tag{5-14}$$

$$\Omega^d_1 = -\frac{1}{R}\frac{1}{2\sqrt{5}}\left(1 + \frac{R\mathcal{G}'}{\mathcal{G}}\right), \tag{5-15}$$

$$\Omega^d_2 = \frac{1}{R\left(c_1 + \frac{2}{\sqrt{5}}\right)^2} f(d_1, c_1), \tag{5-16}$$

$$\Omega^d_3 = \frac{2}{R}\left(\frac{\mathcal{F}}{\mathcal{G}} - 1\right)\frac{d_1 + \frac{1}{\sqrt{5}}}{\left(\frac{2}{\sqrt{5}} + c_1\right)^2}, \tag{5-17}$$

$$\Omega^d_4 = \frac{\Omega_{\mathcal{Q}}}{R^3\sqrt{\mathcal{G}}}, \tag{5-18}$$

$$\Omega^c_{\text{lin}} = -\frac{c_1}{2R}\left(1 + \frac{R\mathcal{G}'}{\mathcal{G}}\right), \tag{5-19}$$

$$\Omega^c_1 = -\frac{1}{R\sqrt{5}}\left(1 + R\frac{\mathcal{G}'}{\mathcal{G}}\right). \tag{5-20}$$

Setting now $u := \binom{c_1}{d_1}$, we rewrite the system under consideration as

$$u' = \frac{1}{R}\left(\begin{pmatrix} -\frac{1}{2} & 1 \\ -\frac{5}{2} & 0 \end{pmatrix} - \frac{1}{2}\left(1 + \frac{\mathscr{G}' R}{\mathscr{G}}\right) I_2\right) u + \omega,$$

where ω contains all the terms $\Omega_i^{c,d}$ apart from Ω_{lin}^d and Ω_{lin}^c and I_2 is the identity matrix. Consider the matrix

$$A = \begin{pmatrix} -\frac{1}{2} & 1 \\ -\frac{5}{2} & 0 \end{pmatrix}$$

and also let

$$B = -\frac{1}{2}\left(1 + \frac{\mathscr{G}' R}{\mathscr{G}}\right) I_2.$$

Then, we find

$$u = \exp \int_{R_0}^{R} \frac{A+B}{R'}\, dR'\, u(R_0) + \int_{R_0}^{R} \left[\exp \int_{R'}^{R} \frac{A+B}{R''}\, dR''\right] \omega(R')\, dR'. \tag{5-21}$$

Note next that

$$\exp \int_{R_0}^{R} \frac{A+B}{R'}\, dR' = \exp \int_{R_0}^{R} \frac{A}{R'}\, dR' \exp \int_{R_0}^{R} \frac{B}{R'}\, dR' = \exp \int_{R_0}^{R} \frac{A}{R'}\, dR' \left(\frac{R_0 \mathscr{G}(R_0)}{R \mathscr{G}(R)}\right)^{\frac{1}{2}}$$

and that the eigenvalues of A are $\lambda_\pm = -\frac{1}{4} \pm \frac{i\sqrt{39}}{4}$. Hence,

$$\left\| \exp \int_{R_0}^{R} \frac{A}{R'}\, dR' \right\| \le C_A \left(\frac{R_0}{R}\right)^{\frac{1}{4}}$$

for some constant $C_A > 0$ depending on the matrix A and we have the following result:

Proposition 5.1. *Provided the corrected energy $\mathscr{G}$ is positive for all $R \in [R_0, R_1]$, one has, for all $R \in [R_0, R_1]$,*

$$|u(R)| \le C_A \left(\frac{R_0}{R}\right)^{\frac{1}{4}} \left(\frac{R_0 \mathscr{G}(R_0)}{R \mathscr{G}(R)}\right)^{\frac{1}{2}} |u(R_0)| + \int_{R_0}^{R} C_A \left(\frac{R'}{R}\right)^{\frac{1}{4}} \left(\frac{R' \mathscr{G}(R')}{R \mathscr{G}(R)}\right)^{\frac{1}{2}} |\omega(R')|\, dR', \tag{5-22}$$

where

$$|\omega| \le C(|\Omega_1^d| + |\Omega_2^d| + |\Omega_3^d| + |\Omega_4^d| + |\Omega_1^c|).$$

It remains to combine the above inequality with our differential inequalities for $\mathscr{G}$ and estimates on the error terms.

5B. ***Source terms of the dynamical system.*** We now combine our results in the latter two sections and we estimate the source terms of the dynamical system. We will assume here that $\mathscr{G}$ is strictly positive, a property that we shall retrieve below in a bootstrap argument.

Estimate for $|\Omega_{\mathcal{Q}}/(R^3\sqrt{\mathcal{G}})|$. Since we have

$$\mathcal{Q} = d R^3 \sqrt{\mathcal{G}} \quad \text{and} \quad \mathcal{Q}_R = d_R R^3 \sqrt{\mathcal{G}} + 3\, d R^2 \sqrt{\mathcal{G}} + d \frac{R^3}{2} \frac{\mathcal{G}'}{\sqrt{\mathcal{G}}},$$

it follows that

$$\left| \frac{\Omega_{\mathcal{Q}}}{R^3 \sqrt{\mathcal{G}}} \right| \leq 2\mathcal{E} R\, d_R + 6 R\mathcal{E} \frac{d}{R} + R\mathcal{E} d \frac{\mathcal{G}'}{\mathcal{G}}.$$

Observe that, while some terms in the right-hand side have no sign, their sum does (because $\mathcal{Q}_R$ is positive).

Estimating $R\mathcal{G}'/\mathcal{G}+1$. From the corrected energy estimate, we get

$$\left| \frac{\mathcal{G}'}{\mathcal{P}} + \frac{\mathcal{G}}{\mathcal{P} R} \right| \leq \frac{\mathcal{P}}{4\pi R^2} \mathcal{E} + \frac{\mathcal{A}}{R^3 \mathcal{P}(R)} + \mathcal{E} \frac{K^2}{2R^3} \left(\int_{S^1} \frac{e^{2\eta}}{a} \right)_R + \frac{\mathcal{P}_R}{R} \mathcal{E};$$

hence,

$$\left| \frac{R\mathcal{G}'}{\mathcal{G}} + 1 \right| \leq \frac{\mathcal{P}}{4\pi R} \frac{\mathcal{F}}{\mathcal{G}} + \frac{\mathcal{A}}{\mathcal{G} R^2} + \frac{\mathcal{F}}{\mathcal{G}} \frac{K^2}{2R^2} \left(\int_{S^1} \frac{e^{2\eta}}{a} \right)_R + \frac{\mathcal{F}}{\mathcal{G}} \mathcal{P}_R \leq \frac{\mathcal{A}}{\mathcal{G} R^2} + \frac{\mathcal{F}}{\mathcal{G}} \frac{\mathcal{Q}_R}{R^2} + \frac{\mathcal{F}}{\mathcal{G}} \frac{\sqrt{\mathcal{G}}}{4\pi} c + \frac{\mathcal{F}}{\mathcal{G}} \sqrt{\mathcal{G}} d. \tag{5-23}$$

Estimates for Ω_1^i. It follows from the estimate (5-23) and the definition of Ω_1^c and Ω_1^d that there exists a constant $C > 0$ such that, for $i = d, c$,

$$|\Omega_1^i| \leq \frac{C}{R} \left(\frac{\mathcal{A}}{\mathcal{G} R^2} + \frac{\mathcal{F}}{\mathcal{G}} \frac{\mathcal{Q}_R}{R^2} + \frac{\mathcal{F}}{\mathcal{G}} \frac{\sqrt{\mathcal{G}}}{4\pi} c + \frac{\mathcal{F}}{\mathcal{G}} \sqrt{\mathcal{G}} d \right). \tag{5-24}$$

Estimates for $\mathcal{F}\mathcal{G}^{-1}$ *and* Ω_3^d. Using Lemma 4.1, we have

$$\left| \frac{\mathcal{F}}{\mathcal{G}} - 1 \right| = \left| \frac{\mathcal{F} - \mathcal{G}}{\mathcal{G}} \right| = \left| \frac{\mathcal{P}\mathcal{G}^U}{\mathcal{G}} \right| \leq \frac{1}{4\pi R} \frac{\mathcal{P}^2 \mathcal{E}}{\mathcal{G}} \leq \frac{\mathcal{P}}{4\pi R} \frac{\mathcal{F}}{\mathcal{G}}. \tag{5-25}$$

As a consequence, provided that c_1 is sufficiently small — so that $\frac{2}{\sqrt{5}} + c_1$ is bounded from below by, say, $\frac{1}{\sqrt{5}}$ — we find

$$|\Omega_3^d| \leq C(|d_1| + 1) \frac{\mathcal{P}}{4\pi R^2} \frac{\mathcal{F}}{\mathcal{G}} \tag{5-26}$$

for some constant $C > 0$.

Note that at this point, we have estimates on all the error terms arising in (5-14)–(5-20), apart from Ω_2^d which will be estimated directly in the next section (using a smallness assumption on c, d).

Estimates on $\mathcal{G}$. After integration of the corrected energy estimate, we find

$$|R\mathcal{G} - R_0 \mathcal{G}(R_0)| \leq \mathcal{A} \left(\frac{1}{R_0} - \frac{1}{R} \right) + \int_{R_0}^{R} \left(\frac{\mathcal{P}}{4\pi R'} \mathcal{F} + \frac{\mathcal{F}\mathcal{Q}_R}{R'^2} + \frac{\mathcal{F}\mathcal{Q}}{R'^3} \right) dR'. \tag{5-27}$$

The last term can be rewritten in terms of $\mathcal{P}_R$, giving

$$|R\mathcal{G} - R_0 \mathcal{G}(R_0)| \leq \mathcal{A} \left(\frac{1}{R_0} - \frac{1}{R} \right) + \int_{R_0}^{R} \left(\frac{\mathcal{P}}{4\pi R'} \mathcal{F} + \frac{\mathcal{F}\mathcal{Q}_R}{R'^2} + \mathcal{F}\mathcal{P}_R \right) dR'. \tag{5-28}$$

6. Small data theory

6A. ***Assumption on the initial data.*** We now restrict ourselves to small data in the following sense. Fix $C_1 > 0$, $\mathscr{A} \in [0, +\infty)$ and $R_0 > 0$, as well as some $\epsilon > 0$. Consider the class of initial data satisfying

$$R_0 \mathscr{G}(R_0) - \frac{\mathscr{A}}{R_0} \geq C_1 > 0, \tag{6-1}$$

$$|c_1|(R_0) \leq \epsilon, \tag{6-2}$$

$$|d_1|(R_0) \leq \epsilon, \tag{6-3}$$

$$\left|\frac{\mathscr{F}}{\mathscr{G}} - 1\right|(R_0) \leq 1, \tag{6-4}$$

$$\mathscr{G}(R_0) + \frac{\mathscr{A}}{R_0^2} \leq \epsilon, \tag{6-5}$$

where $\mathscr{A} = R_0^2 \left(\int_{S^1} a^{-1} U_R\right)^2 (R_0)$.

Note that the first assumption implies in particular that $\mathscr{G} > 0$. The second and third assumptions imply that $\mathscr{P}$ and $\mathscr{P}_R$ are close to their expected asymptotic behavior (which depends on $\mathscr{E}$, hence the need for normalized quantities). The fourth condition implies that the correction term $\mathscr{G}^U$ is "not too large" compared to the energy $\mathscr{E}$. The last inequality means that the (rescaled) energy is small.

Let R_b be the largest time R such that the following bootstrap assumptions are valid in $\mathscr{B} := [R_0, R_b)$. For all $R \in \mathscr{B}$, we have

$$|c_1|(R) < \epsilon^{1/4}, \tag{6-6}$$

$$|d_1|(R) < \epsilon^{1/4}, \tag{6-7}$$

$$\left|\frac{\mathscr{F}}{\mathscr{G}} - 1\right|(R) < 2, \tag{6-8}$$

$$0 < \mathscr{G}(R_0) < \left(R_0 \mathscr{G}(R_0) + \frac{\mathscr{A}}{R_0}\right)\frac{2}{R}. \tag{6-9}$$

The set $\mathscr{B}$ is clearly open in $[R_0, +\infty)$. Moreover, $\mathscr{B}$ is nonempty, by the smallness assumptions.

As an immediate consequence of (6-6) and (6-7), if ϵ is sufficiently small then we have, in $\mathscr{B}$,

$$\frac{1}{c^2}(R_0) = \frac{1}{\left(c_1 + \frac{2}{\sqrt{5}}\right)^2}(R_0) \leq 2, \tag{6-10}$$

$$|c| = \left|\frac{2}{\sqrt{5}} + c_1\right| \leq 1, \tag{6-11}$$

$$|d| = \left|\frac{1}{\sqrt{5}} + d_1\right| \leq 1. \tag{6-12}$$

Furthermore, from (6-9) and (6-5), we have immediately, in $\mathscr{B}$,

$$\mathscr{G} \leq 2\frac{R_0}{R}\left(\mathscr{G}(R_0) + \frac{\mathscr{A}}{R_0^2}\right) \leq 2\epsilon \frac{R_0}{R} \leq 2\epsilon. \tag{6-13}$$

We now consider C_1 and $\mathscr{A}$ as fixed in (6-1). We will show that there exists an $\epsilon_0 > 0$ and a constant $r > 0$ such that, for all $0 < \epsilon < \epsilon_0$ and $R_0 > r$, the set $\mathscr{B}$ is closed; this will be done by "improving" each of the bootstrap assumptions (6-6)–(6-9). Moreover, ϵ_0 will depend only on a lower bound for r (as well as $\mathscr{A}$ and C_1).

6B. ***Improving the assumption on*** $\mathscr{F}\mathscr{G}^{-1}$**.** In view of the estimate (5-25), we have

$$\left|\frac{\mathscr{F}}{\mathscr{G}} - 1\right| \leq 3\frac{c\mathscr{G}^{1/2}}{4\pi} \leq \frac{3\sqrt{2}}{4\pi}\epsilon^{1/2}, \tag{6-14}$$

by using the bootstrap assumptions (6-8) and (6-11), and using (6-13). This improves (6-8).

Throughout, the letter C will be used to denote numerical constants that are independent of ϵ and R_0 and may change at each occurrence. Thus, the above estimate reads

$$\left|\frac{\mathscr{F}}{\mathscr{G}} - 1\right| \leq C\epsilon^{1/2}.$$

Improving the $\mathscr{G}$ *assumption.* From the corrected energy estimate (5-28), we have

$$R\mathscr{G} \leq R_0\mathscr{G}(R_0) + \frac{\mathscr{A}}{R_0} + \int_{R_0}^{R} R'\mathscr{G}\frac{\mathscr{F}}{\mathscr{G}}\left(\frac{\mathscr{P}}{4\pi R'^2} + \frac{\mathscr{Q}_R}{R'^3} + \frac{\mathscr{P}_R}{R'}\right) dR';$$

hence,

$$\mathscr{G} \leq \frac{D_0}{R} \exp \int_{R_0}^{R} (1 + C\epsilon^{1/2})\left[\frac{\mathscr{P}}{4\pi R'^2} + \left(\mathscr{Q}_R R'^{-3} + \frac{\mathscr{P}_R}{R'}\right)\right],$$

where $D_0 = R_0\mathscr{G}(R_0) + \mathscr{A}/R_0$ and we have used the improved inequality (6-14).

The integral $\int_{R_0}^{R}(\mathscr{P}/(4\pi R'^2))\,dR'$ can be estimated using (6-13):

$$\int_{R_0}^{R} \frac{\mathscr{P}}{4\pi R'^2}\,dR' = \int_{R_0}^{R} \frac{cR'\mathscr{G}^{1/2}}{4\pi R'^2}\,dR' \leq \int_{R_0}^{R} \frac{C\epsilon^{1/2}R_0^{1/2}}{4\pi R'^{3/2}}\,dR' \leq C\epsilon^{1/2}$$

for some fixed numerical constant $C > 0$.

For the other integrals, we integrate by parts:

$$\begin{aligned}\int_{R_0}^{R}\left(\frac{\mathscr{Q}_R}{R'^3} + \frac{\mathscr{P}_R}{R'}\right) dR' &\leq \frac{\mathscr{Q}}{R^3} + \frac{\mathscr{P}}{R} + \int_{R_0}^{R}\left[\frac{3\mathscr{Q}}{R'^4} + \frac{\mathscr{P}}{R'^2}\right] dR' \\ &\leq (c+d)\mathscr{G}^{1/2} + \int_{R_0}^{R} \frac{3d+c}{R'}\mathscr{G}^{1/2}(R')\,dR' \\ &\leq C\epsilon^{1/2} + C\int_{R_0}^{R} \frac{R_0^{1/2}}{R'^{3/2}}\epsilon^{1/2}\,dR' \\ &\leq C\epsilon^{1/2}.\end{aligned}$$

Combining this result with the previous estimate, we have thus obtained

$$R\mathscr{G} \leq D_0 \exp((1 + C\epsilon^{1/2})C\epsilon^{1/2}) < \tfrac{3}{2}D_0 \tag{6-15}$$

providing that ϵ is small enough. This improves (6-9).

A lower bound on $\mathscr{G}$. We derive here a lower bound on $R\mathscr{G}$. From the corrected energy inequality in differential form (4-5) and the estimates on the error term, we have

$$\frac{d}{dR}(R\mathscr{G}) \geq -\frac{\mathscr{A}}{R^2} - R\mathscr{G}\left[\frac{\mathscr{F}}{\mathscr{G}}\left(\frac{\mathscr{P}}{4\pi R'^2} + \frac{\mathscr{Q}_R}{R'^3} + \frac{\mathscr{P}_R}{R'}\right)\right]. \tag{6-16}$$

Let

$$\Omega' = \frac{\mathscr{F}}{\mathscr{G}}\left(\frac{\mathscr{P}}{4\pi R'^2} + \frac{\mathscr{Q}_R}{R'^3} + \frac{\mathscr{P}_R}{R'}\right).$$

The estimates of the previous sections have shown that

$$\int_{R_0}^{R} \Omega'\, dR' \leq C\epsilon^{1/2}.$$

We can rewrite (6-16) as

$$\frac{d}{dR}(R\mathscr{G}) \geq -\frac{\mathscr{A}}{R^2} - R\mathscr{G}\Omega',$$

leading to

$$\begin{aligned}\frac{d}{dR}\left(R\mathscr{G}\exp\int_{R_0}^{R}\Omega'\,dR'\right) &\geq -\frac{\mathscr{A}}{R^2}\exp\int_{R_0}^{R}\Omega'\,dR' \\ &= \frac{d}{dR}\left(\frac{\mathscr{A}}{R}\right)\exp\int_{R_0}^{R}\Omega'\,dR', \\ &= \frac{d}{dR}\left(\frac{\mathscr{A}}{R}\exp\int_{R_0}^{R}\Omega'\,dR'\right) - \frac{\mathscr{A}}{R}\Omega'\exp\int_{R_0}^{R}\Omega'\,dR'.\end{aligned}$$

Thus,

$$\frac{d}{dR}\left[\left(R\mathscr{G} - \frac{\mathscr{A}}{R}\right)\exp\int_{R_0}^{R}\Omega'\,dR'\right] \geq -\frac{\mathscr{A}}{R}\Omega'\exp\int_{R_0}^{R}\Omega'\,dR',$$

which leads after integration to

$$R\mathscr{G} - \frac{\mathscr{A}}{R} \geq \left(R_0\mathscr{G}(R_0) - \frac{\mathscr{A}}{R_0}\right)(1 - C\epsilon^{1/2}) - \frac{\mathscr{A}}{R_0}C\epsilon^{1/2} = C_1(1 - C\epsilon^{1/2}) - \frac{\mathscr{A}}{R_0}C\epsilon^{1/2} \geq \frac{C_1}{2} \tag{6-17}$$

provided that ϵ is sufficiently small, depending on $\mathscr{A}$, C_1 and a lower bound on R_0.

Since $\mathscr{A} \geq 0$, we have thus obtained $R\mathscr{G} \geq \frac{1}{2}C_1$. In particular, we have improved the lower bound bootstrap inequality for $\mathscr{G}$.

Remark 6.1. Instead of starting from the corrected energy inequality in differential form, one could use here the estimate (5-28) as well as the estimates of the previous section to estimate the term containing $\mathscr{G}$ in the error term. This would lead to an estimate of the form

$$R\mathscr{G} \geq C_1 - D_0 C\epsilon^{1/2}$$

and would therefore require ϵ to be small compared to D_0. The above method has the advantage of not constraining ϵ any further.

Improving the c_1 and d_1 assumptions. Using the lower bound on $\mathscr{G}$ just obtained, the bootstrap assumption (6-9), the initial data assumptions (6-2) and (6-3) and the fact that $R'/R \leq 1$ if $R' \in [R_0, R]$, it follows from (5-22) that

$$|u| \leq \left(\frac{C_A D_0}{C_1}\right)^{\frac{1}{2}} \epsilon + C\left(\frac{4D_0}{C_1}\right)^{\frac{1}{2}} \int_{R_0}^{R} (|\Omega_1^c| + |\Omega_1^d| + |\Omega_4^d|)\, dR' + C\left(\frac{4D_0}{C_1}\right)^{\frac{1}{2}} \int_{R_0}^{R} \left(\frac{R'}{R}\right)^{\frac{1}{4}} (|\Omega_2^d| + |\Omega_3^d|)\, dR'. \tag{6-18}$$

We now estimate all the error terms in ω. First, we have

$$|\Omega_1^c, \Omega_1^d| \leq \frac{C}{R}\left|1 + R\frac{\mathscr{G}'}{\mathscr{G}}\right| \leq \frac{C}{R}\left(\frac{2}{C_1}\frac{\mathscr{A}}{R} + C\frac{\mathscr{Q}_R}{R^2} + C\mathscr{G}^{1/2}\right), \tag{6-19}$$

using (5-23), (5-24) and (6-8). The first term in the parentheses in the right-hand side of the last inequality will contribute to (6-18) as

$$\left(\frac{4D_0}{C_1}\right)^{\frac{1}{2}} \int_{R_0}^{R} \frac{2}{C_1}\frac{\mathscr{A}}{R'^2}\, dR' \leq C\frac{\mathscr{A}}{C_1^{3/2}} D_0^{1/2} R_0^{-1} \leq C\frac{\mathscr{A}}{C_1^{3/2} R_0^{1/2}} (D_0 R_0^{-1})^{1/2} \leq C(C_1, R_0, \mathscr{A})\epsilon^{1/2},$$

by using the smallness assumption (6-5). The second term can be estimated using an integration by parts, leading to the estimate

$$C\left(\frac{D_0}{C_1}\right)^{\frac{1}{2}} \int_{R_0}^{R} \frac{\mathscr{Q}_R}{R'^3}\, dR' \leq C\left(\frac{D_0}{C_1}\right)^{\frac{1}{2}} \epsilon^{1/2}.$$

Since $D_0/C_1 = 1 - 2\mathscr{A}/(C_1 R_0)$, we thus obtain

$$C\left(\frac{D_0}{C_1}\right)^{\frac{1}{2}} \int_{R_0}^{R} \frac{\mathscr{Q}_R}{R'^3}\, dR' \leq C\epsilon^{1/2},$$

by choosing ϵ sufficiently small, depending only on a lower bound on C_1 and $\mathscr{A}$ and a lower bound on R_0.

The last term in (6-19) can be estimated using (6-13) leading to

$$\int_{R_0}^{R} \frac{\mathscr{G}^{1/2}}{R'}\, dR' \leq C\epsilon^{1/2}.$$

The estimates for Ω_2^d and Ω_3^d are straightforward using the bootstrap assumptions

$$|\Omega_2^d| \leq \frac{C}{R}\epsilon^{1/2} \quad \text{and} \quad |\Omega_3^d| \leq \frac{C}{R}\epsilon^{1/2}.$$

For Ω_4^d, we note that, in view of (5-18) and (5-3), we have

$$|\Omega_4^d| \leq \frac{2R\mathscr{E}\mathscr{Q}_R}{R^3\sqrt{\mathscr{G}}}.$$

Then, we note that

$$\mathscr{E} = \frac{\mathscr{F}}{\mathscr{P}} = \frac{\mathscr{F}}{cR\mathscr{G}^{1/2}};$$

hence,

$$\mathscr{E}\mathscr{G}^{-1/2} = \frac{1}{cR}\left(\frac{\mathscr{F}}{\mathscr{G}}\right).$$

Using the bootstrap assumptions, this leads to

$$|\Omega_4^d| \leq \frac{1}{c}\frac{\mathcal{F}}{\mathcal{G}}\frac{2\mathcal{Q}_R}{R^3} \leq C\mathcal{Q}_R R^{-3}, \tag{6-20}$$

where we have used that $\mathcal{Q}_R \geq 0$ in the last estimate. Its integral can then be estimated by integration by parts, as we have already done previously.

Combining all these estimates leads us to

$$|u| \leq C\left(\frac{2D_0}{C_1}\right)^{\frac{1}{2}}\epsilon + C(C_1, R_0, \mathcal{A})\epsilon^{1/2} \leq C(\mathcal{A}, R_0, C_1)\epsilon^{1/2},$$

which improves (6-6) and (6-7). In conclusion, we have improved all of the bootstrap inequalities and it follows that

$$\mathcal{B} = [R_0, +\infty).$$

7. The asymptotic regime

In this section, we state and prove our main result.

Theorem 7.1 (late-time asymptotics of T^2-symmetric polarized vacuum spacetimes). *Let $\mathcal{A} \geq 0$ and let $C_1 > 0$ and $r > 0$ be fixed constants. Then there exists an ϵ_0 such that, if $0 \leq \epsilon \leq \epsilon_0$ and $R_0 \geq r$ then, for any initial data set satisfying the smallness conditions* (6-1)–(6-5), *the associated solution has the following asymptotic behavior*: *for all times $R \geq R_0$ and all $\theta \in S^1$,*

$$|u|(R,\theta) = O(R^{-1/4}), \tag{7-1}$$

$$|R\mathcal{G}(R) - C_\infty| = O(R^{-1/2}), \tag{7-2}$$

$$\left|\mathcal{P}(R) - \frac{2}{\sqrt{5}}C_\infty^{1/2}R^{1/2}\right| = O(R^{1/4}), \tag{7-3}$$

$$\left|\mathcal{Q}(R) - \frac{1}{\sqrt{5}}C_\infty^{1/2}R^{5/2}\right| = O(R^{9/4}), \tag{7-4}$$

$$\left|\mathcal{E}(R) - \frac{\sqrt{5}C_\infty^{1/2}}{2R^{3/2}}\right| = O(R^{-7/4}), \tag{7-5}$$

$$\left|\frac{1}{2\pi}\int_{S^1}\eta(R,\theta')\,d\theta' - \eta(R,\theta)\right| = O(R^{-1/2}), \tag{7-6}$$

$$|K^2e^{2\eta}(R,\theta) - R^2| = O(R^{7/4}), \tag{7-7}$$

$$|a^{-1}(R,\theta)\mathcal{P}^{-1}(R) - \mathcal{L}(\theta)| = O(R^{-1/2}), \tag{7-8}$$

$$\left|\frac{1}{2\pi}\int_{S^1}U(R,\theta)\,d\theta - U(R,\theta)\right| = O(R^{-1/2}), \tag{7-9}$$

$$|U(R,\theta) - C_U| = O(R^{-1/2}), \tag{7-10}$$

$$\left|H(R,\theta) - \frac{4}{K\sqrt{5}}C_\infty^{1/2}R^{1/2}\mathcal{L}(\theta)\right| = O(R^{1/4}), \tag{7-11}$$

where $C_\infty > 0$ and C_U are constants depending on the solution and $\mathscr{L}(\theta)$ is a $W^{1,1}(S^1)$ strictly positive function.

Proof. Most of the above estimates are simply obtained by revisiting the proof in the previous section and checking that the error terms are now *integrable*.

For instance, in order to prove (7-1), note that, from (5-22) and the estimates of Section 6, we have

$$|u| \le C R^{-1/4}\biggl(1+\int_{R_0}^{R} R'^{1/4}|\omega(R')|\,dR'\biggr). \tag{7-12}$$

From (6-19) and (6-20), one can easily see that the contributions of Ω_1^c, Ω_1^d and Ω_4^d are integrable in R. For instance, using an integration by parts,

$$\begin{aligned}\int_{R_0}^{R} \frac{\mathfrak{Q}_R}{R'^{3-1/4}} &\le C\frac{\mathfrak{Q}}{R^{3-1/4}} + C\int_{R_0}^{R}\frac{\mathfrak{Q}}{R'^{4-1/4}}\,dR',\\ &\le C\frac{\mathfrak{Q}}{R^3\mathscr{G}^{1/2}}(R\mathscr{G})^{1/2}R^{-1/4} + C\int_{R_0}^{R}\frac{\mathfrak{Q}}{R'^3\mathscr{G}^{1/2}}(R'\mathscr{G})^{1/2}R'^{-5/4}\,dR',\\ &\le CR^{-1/4} + C\int_{R_0}^{R} R'^{-5/4}\,dR'\\ &\le C.\end{aligned}$$

For Ω_3^d, it follows from (5-26) and the estimates of the previous section that $|\Omega_3^d| \le CR^{-3/2}$. Thus, its contribution to the integral of (7-12) is integrable. Since, moreover, $|\Omega_2^d| \le (C/R)|u|^2$, (7-12) has now been reduced to

$$|u| \le C R^{-1/4}\biggl(1+\int_{R_0}^{R} R'^{-3/4}|u|^2(R')\,dR'\biggr). \tag{7-13}$$

Since we already know from the estimates of the previous section that $|u| \le C\epsilon^{1/2}$, an application of Gronwall's lemma gives us the weak bound

$$R^{1/4}|u| \le CR^{\epsilon^{1/2}}.$$

It then follows that $R^{-3/4}|u|^2 \le CR^{-5/4+\epsilon}$ and thus, for ϵ sufficiently small, (7-13) now implies the desired estimate (7-1).

Similarly, to prove (7-2), first note that $d(R\mathscr{G})/dR$ is integrable, using the estimates of Section 6 and (5-23). Thus, there exists a constant C_∞ such that $R\mathscr{G} \to C_\infty$ as $R \to +\infty$. Since $R\mathscr{G}$ is uniformly bounded from below in view of (6-17), we have $C_\infty > 0$. To get the rate of convergence, it then suffices to write $R\mathscr{G} - C_\infty = \int_R^\infty (d(R'\mathscr{G})/dR)\,dR'$ and to estimate the integral as before.

Then, (7-3), (7-4) and (7-5) follow from the definitions of $\mathscr{P}$, $\mathfrak{Q}$ and $\mathscr{E}$.

For (7-6), using (2-7), the simple estimate $F \le E$ and (7-5), we have, for all $R \ge R_0$ and $\theta \in S^1$,

$$\biggl|\frac{1}{2\pi}\int_{S^1}\eta(R,\theta')\,d\theta' - \eta(R,\theta)\biggr| \le \int_{S^1}|\eta_\theta|(R,\theta')\,d\theta' \le \int_{S^1} RF\,d\theta' \le \int_{S^1} RE \le CR^{-1/2}$$

for some $C > 0$. For (7-7), we use (7-6), (7-3) and (7-4) as well as

$$\begin{aligned}\mathscr{P}\tfrac{1}{2}K^2e^{2\eta}(R,\theta) &= \int_{S^1} a^{-1}(R,\theta')\tfrac{1}{2}K^2e^{2(\eta(R,\theta)-\eta(R,\theta')+\eta(R,\theta'))}\,d\theta' \\ &= \int_{S^1} a^{-1}(R,\theta')\tfrac{1}{2}K^2e^{2(\eta(R,\theta')+O(R^{-1/2}))}\,d\theta' \\ &= \mathscr{Q}(1+O(R^{-1/2})).\end{aligned}$$

For (7-8), we first differentiate (2-5) in θ; that is,

$$(2\ln a)_{R\theta} = -\frac{K^2}{R^3}e^{2\eta}2\eta_\theta. \tag{7-14}$$

Note that the right-hand side is integrable in $L([R_0,+\infty)\times S^1)$ since

$$\int_{R_0}^{\infty}\int_{S^1}\left|\frac{K^2}{R^3}e^{2\eta}2\eta_\theta\right|d\theta\,dR \le \int_{R_0}^{R} CR^{-1}R\mathscr{E} \le C, \tag{7-15}$$

in view of (7-5). This implies that $(\ln a)_\theta(R,\theta)$ converges in $L^1(S^1)$ as $R\to+\infty$ to some function $\mathscr{R}(\theta)\in L^1(S^1)$ and, moreover, we have the estimate

$$\|(\ln a)_\theta - \mathscr{R}\|_{L^1(S^1)} = O(R^{-1/2}),$$

by using (7-15).

Integrating over $[\theta,\theta']$, we get

$$\frac{a(R,\theta)}{a(R,\theta')} = \exp\biggl(\int_{\theta'}^{\theta}\mathscr{R}(\theta'')\,d\theta'' + O(R^{-1/2})\biggr).$$

Integrating again in the θ' variable, we get

$$\left|a(R,\theta)\mathscr{P} - \int_{S^1} e^{\int_{\theta'}^{\theta}\mathscr{R}(\theta'')\,d\theta''}\,d\theta'\right| \le C\bigl(\exp\,(O(R^{-1/2}))-1\bigr) = O(R^{-1/2}).$$

For (7-11), it is sufficient to note that, with the knowledge of the asymptotic behavior of a and η, and (2-8), we can integrate H_R directly and then compute the integral up to some error.

The property (7-9) is an easy consequence of (7-8), (7-3) and (7-5). For (7-10), we observe that

$$\left|\frac{d}{dR}\int_0^{2\pi} U\,d\theta\right| = \left|\int_0^{2\pi} U_R\,d\theta\right| \le (2\pi)^{1/2}\biggl(\int_0^{2\pi} U_R^2\,d\theta\biggr)^{\frac{1}{2}}$$

and

$$\begin{aligned}\biggl(\int_0^{2\pi} U_R^2\,d\theta\biggr)(R) = \biggl(\int_0^{2\pi} a^{-1}aU_R^2\,d\theta\biggr)(R) &\le \sup_{[0,2\pi]} a(R,\theta)\int_0^{2\pi} a^{-1}U_R^2\,d\theta \\ &\le \biggl(\frac{1}{\mathscr{P}}+o(a)\biggr)\frac{1}{\mathscr{L}(\theta)}\int_0^{2\pi} a^{-1}U_R^2\,d\theta \\ &\le \frac{C}{R^2}\end{aligned}$$

for some $C > 0$. Here we have used (7-8) together with the fact $\mathscr{L}$ is bounded away from zero uniformly, as well as (7-3) and (7-5).

This implies that

$$\left| \frac{d}{dR} \int_0^{2\pi} U \, d\theta \right| \le \frac{C}{R}$$

and, by integration and (7-9), we obtain the rough bound on U,

$$|U| \le C \ln R.$$

Applying now the commutator estimate from Lemma 3.2, we have that

$$\left| \frac{d}{dR} \langle \widetilde{U} \rangle - \langle \widetilde{U}_R \rangle \right| \le \frac{\pi K^2}{R^3} \langle |\widetilde{U}| \rangle \| (e^{2\eta})_\theta \|_{L^1(S^1)}. \tag{7-16}$$

From the above rough bound on U, we have

$$\left| \langle |\widetilde{U}| \rangle \right| \le C \ln R.$$

Moreover, one can estimate $\|(e^{2\eta})_\theta\|_{L^1(S^1)}$ as before to get

$$\|(e^{2\eta})_\theta\|_{L^1(S^1)} \le C R^{3/2}.$$

Thus, the right-hand side of (7-16) is integrable in R. Since, moreover,

$$\langle \widetilde{U}_R \rangle = \frac{1}{\mathscr{P}} \int_0^{2\pi} U_R a^{-1}(R, \theta) \, d\theta = \frac{R_0}{\mathscr{P} R} \int_0^{2\pi} U_R a^{-1}(R_0, \theta) \, d\theta$$

using the conservation law in Lemma 3.3, it follows that $\langle \widetilde{U}_R \rangle$ and, therefore, $d\langle \widetilde{U} \rangle / dR$ are integrable. By having checked the convergence of all the integrals involved in our analysis, this completes the proof of (7-10) and, thus, of Theorem 7.1. □

8. Future geodesic completeness

In this section, we complete the proof of the geodesic completeness property under the smallness assumption (6-1)–(6-5). There are only small modifications in comparison to the proof already presented by the authors in [LeFloch and Smulevici 2016] for weakly regular Gowdy spacetimes. One of difficulties (observed and solved in [loc. cit.]) is that, with limited control of the Christoffel symbols in the L^1 or L^2 norms (in space) only, the local existence of geodesics is not guaranteed by the standard Cauchy–Lipschitz theorem. Instead, we first established that the Christoffel symbols admit traces along timelike curves and we relied on a compactness argument à la Arzelá–Ascoli in order to establish the existence of geodesics. This part of the analysis can be repeated here almost identically in our T^2 setting, by using the estimates in [LeFloch and Smulevici 2015] for the compactness argument. (This compactness is required in the proof of existence of traces, as explained in Proposition 3.5 of [LeFloch and Smulevici 2016]). We do not repeat these arguments here and directly assume the existence of geodesics (which, for instance, is immediate in the smooth case).

Theorem 8.1 (future geodesic completeness). *Let $(\mathcal{M}, g)$ be a nonflat, polarized, T^2-symmetric, vacuum spacetime with weak regularity whose initial data set satisfies the conditions* (6-1)–(6-5). *Then all future timelike geodesics are future complete.*

Proof. For simplicity in the presentation, we focus on the smooth case. Let ξ be a future maximal timelike geodesic defined on an interval $[s_0, s_1)$. We have $g(\dot{\xi}, \dot{\xi}) < 0$ and

$$\ddot{\xi}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{\xi}^\beta\dot{\xi}^\gamma = 0. \tag{8-1}$$

Following [LeFloch and Smulevici 2016], we observe that, since X and Y are Killing fields, $J_X = g(\dot{\xi}, X)$ and $J_Y = g(\dot{\xi}, Y)$ are constant along ξ, so that $J_X = e^{2U}(\dot{\xi}^X + G\dot{\xi}^\theta)$ and $J_Y = e^{-2U}R^2(\dot{\xi}^Y + H\dot{\xi}^\theta)$ are constants along ξ. We use the same strategy as in [ibid., Section 4]. First, by standard arguments (see [ibid., Lemma 4.10]), it follows that $R(\xi(s)) \to +\infty$ as $s \to s_1$. Then, since $R(\xi(s)) - R(\xi(s_0)) = \int_{s_0}^s \dot{\xi}^R\, ds$, it follows that any bound of the form $\dot{\xi}^R < CR^p$ for $p < 1$ implies that $s_1 = +\infty$. Note also that, since $R(\xi(s)) \to +\infty$, given any $R' > 0$ we may assume, without loss of generality, that $R(\xi(s_0)) \geq R'$.

We now analyze the structure of the equation satisfied by $\dot{\xi}^R$,

$$\ddot{\xi}^R + \Gamma^R_{\beta\gamma}\dot{\xi}^\beta\dot{\xi}^\gamma = 0. \tag{8-2}$$

The term $\Gamma^R_{\beta\gamma}\dot{\xi}^\beta\dot{\xi}^\gamma = 0$ is decomposed in the form

$$\Gamma^R_{\beta\gamma}\dot{\xi}^\beta\dot{\xi}^\gamma = \Gamma^R_{RR}\dot{\xi}^R\dot{\xi}^R + \Gamma^R_{\theta\theta}\dot{\xi}^\theta\dot{\xi}^\theta + 2\Gamma^R_{R\theta}\dot{\xi}^R\dot{\xi}^\theta + 2\Gamma^R_{\theta a}\dot{\xi}^\theta\dot{\xi}^a + \Gamma^R_{ab}\dot{\xi}^a\dot{\xi}^b,$$

where $\{a, b\} = \{X, Y\}$. Recall now that

$$\Gamma^R_{RR} = \eta_R - U_R, \tag{8-3}$$

$$\Gamma^R_{\theta\theta} = \frac{\eta_R - U_R}{a^2} - \frac{a_R}{a^3} + e^{2U}U_R G^2 e^{-2(\eta-U)} + (e^{-2U}R^2H^2)_R \frac{e^{-2(\eta-U)}}{2}, \tag{8-4}$$

$$\Gamma^R_{R\theta} = \eta_\theta - U_\theta. \tag{8-5}$$

Observe also that

$$\eta_R - U_R = R\left(\left(U_R - \frac{1}{2R}\right)^2 + a^2U_\theta^2\right) - \frac{1}{4R} - \frac{K^2}{4R^3}e^{2\eta},$$

while

$$\eta_\theta - U_\theta = 2R\left(U_R - \frac{1}{2R}\right)U_\theta.$$

As a consequence, it follows that the following quadratic form inequality holds:

$$(\eta_R - U_R)(dR^2 + a^{-2}\,d\theta^2) + 2(\eta_\theta - U_\theta)\,dR\,d\theta + \left(\frac{1}{4R} + \frac{K^2}{4R^3}e^{2\eta}\right)(dR^2 + a^{-2}\,d\theta^2) \geq 0. \tag{8-6}$$

Returning now to (8-2), this leads us to

$$\begin{aligned}\ddot{\xi}^R \leq \left(\frac{1}{4R} + \frac{K^2}{4R^3}e^{2\eta}\right)((\dot{\xi}^R)^2 + a^{-2}(\dot{\xi}^\theta)^2) + \frac{a_R}{a^3}(\dot{\xi}^\theta)^2 \\ - \left(e^{2U}U_R G^2 e^{-2(\eta-U)} + (e^{-2U}R^2H^2)_R \frac{e^{-2(\eta-U)}}{2}\right)(\dot{\xi}^\theta)^2 - 2\Gamma^R_{\theta a}\dot{\xi}^\theta\dot{\xi}^a - \Gamma^R_{ab}\dot{\xi}^a\dot{\xi}^b.\end{aligned}$$

Note that the term containing a_R/a^3 has the right sign and can absorb the term $K^2 e^{2\eta}(\dot\xi^\theta)^2/(4R^3)$. Using moreover the estimate (7-7) and the fact that $|a^{-1}\dot\xi^\theta| \le \dot\xi^R$, for all $\epsilon > 0$ we may assume that $R\xi(s_0)$ is sufficiently large that

$$\ddot\xi^R \le \left(\frac{3+\epsilon}{4R}\right)(\dot\xi^R)^2 - 2\Gamma^R_{\theta a}\dot\xi^\theta\dot\xi^a - \Gamma^R_{ab}\dot\xi^a\dot\xi^b - \left(e^{2U}U_R G^2 e^{-2(\eta-U)} + (e^{-2U}R^2H^2)_R \frac{e^{-2(\eta-U)}}{2}\right)(\dot\xi^\theta)^2.$$

Recalling now that $dR/ds = \dot\xi^R$, the last inequality can be rewritten as

$$\begin{aligned}\frac{d}{ds}&(R^{-3/4-\epsilon}\dot\xi^R)\\ &\le R^{-3/4-\epsilon}\big(-\big(e^{2U}U_R G^2 e^{-2(\eta-U)} + (e^{-2U}R^2H^2)_R \tfrac12 e^{-2(\eta-U)}\big)(\dot\xi^\theta)^2 - 2\Gamma^R_{\theta a}\dot\xi^\theta\dot\xi^a - \Gamma^R_{ab}\dot\xi^a\dot\xi^b\big).\end{aligned} \tag{8-7}$$

For the three terms in the right-hand side, recall that

$$\begin{aligned}\Gamma^R_{X\theta} &= e^{-2\eta}e^{4U}U_R G,\\ \Gamma^R_{Y\theta} &= \tfrac12 e^{-2(\eta-U)}(e^{-2U}R^2H)_R,\\ \Gamma^R_{XX} &= e^{-2\eta}e^{4U}U_R,\\ \Gamma^R_{XY} &= 0,\\ \Gamma^R_{YY} &= \tfrac12 e^{-2(\eta-U)}(e^{-2U}R^2)_R.\end{aligned}$$

These terms can be combined with the terms containing H^2 and G^2 above arising from $\Gamma^R_{\theta\theta}$ as follows:

$$\Gamma^R_{XX}(\dot\xi^X)^2 + 2\Gamma^R_{\theta X}\dot\xi^\theta\dot\xi^X + e^{2U}U_R G^2 e^{-2(\eta-U)}(\dot\xi^\theta)^2 = e^{-2(\eta-U)}e^{2U}U_R(\dot\xi^X + G\dot\xi^\theta)^2 = e^{-2\eta}U_R\, J_X^2$$

and

$$\begin{aligned}\Gamma^R_{YY}&(\dot\xi^Y)^2 + 2\Gamma^R_{\theta Y}\dot\xi^\theta\dot\xi^Y + (e^{-2U}R^2H^2)_R\tfrac12 e^{-2(\eta-U)}(\dot\xi^\theta)^2\\ &= \tfrac12 e^{-2(\eta-U)}\big((e^{-2U}R^2)_R(\dot\xi^Y + H\dot\xi^\theta)^2 + e^{-2U}R^2 2HH_R(\dot\xi^\theta)^2 + 2e^{-2U}H_R R^2\dot\xi^\theta\dot\xi^Y\big)\\ &= \tfrac12 e^{-2(\eta-U)}((e^{-2U}R^2)_R R^{-4}e^{4U}J_Y^2 + 2H_R\dot\xi^\theta J^Y).\end{aligned}$$

Now let $\mu = \eta - U + \frac14\ln R - \frac12\ln a$. Note that

$$\mu_R = R\left(\left(U_R - \frac{1}{2R}\right)^2 + a^2U_\theta^2\right) \ge 0.$$

Then, using that U is uniformly bounded and (7-7), we easily have the estimates

$$|e^{-2\eta}U_R\, J_X^2| \le CR^{-2}\left(R^{-1/2}\mu_R^{1/2} + \frac1R\right), \tag{8-8}$$

$$\left|\frac{e^{-2(\eta-U)}}{2}(e^{-2U}R^2)_R R^{-4}e^{4U}J_Y^2\right| \le CR^{-4}\left(R^{-1/2}\mu_R^{1/2} + \frac1R\right), \tag{8-9}$$

for some constant $C > 0$. Moreover, in view of (2-8), (7-7) and the estimate $|\dot\xi^\theta| \le a\dot\xi^R$,

$$|e^{-2(\eta-U)}H_R\dot\xi^\theta J^Y| \le C\frac{\dot\xi^R}{R^3}.$$

Returning to (8-7), we obtain

$$\frac{d}{ds}(R^{-3/4-\epsilon}\dot{\xi}^R) \le CR^{-13/4-\epsilon}(\mu_R^{1/2}+R^{-1/2}) + C\frac{\dot{\xi}^R}{R^3}.$$

The second term in the right-hand side is integrable since $\dot{\xi}^R = dR\xi(s)/ds$. Moreover, $R^{-13/4-\epsilon}R^{-1/2}$ is decreasing in R and, therefore, integrable on any bounded interval $[s_0, s_1]$. Thus, it remains only to show that $R^{-13/4-\epsilon}\mu_R^{1/2}$ is integrable.

Let $M^2 = -g(\dot{\xi}, \dot{\xi})$. Then we have

$$a^{-2}\left(\frac{\dot{\xi}^\theta}{\dot{\xi}^R}\right)^2 \le 1 - \frac{M^2 e^{-2(\eta-U)}}{(\dot{\xi}^R)^2}.$$

Let $\chi = M^2 e^{-2(\eta-U)}/(\dot{\xi}^R)^2 \le 1$ and let $\rho = \eta - U$. Then we find[3]

$$\frac{d\rho}{ds} + \frac{1}{4}\frac{d}{ds}(\ln R) - \frac{a_R}{2a}\dot{\xi}^R \ge (1-(1-\chi)^{1/2})\mu_R\dot{\xi}^R \ge \frac{1}{2}\chi\mu_R\dot{\xi}^R. \tag{8-10}$$

In particular, $d\rho/ds + \frac{1}{4}d(\ln R)/ds - (a_R/(2a))\dot{\xi}^R \ge 0$. As a consequence, we have

$$\mu_R \le 2\left(\frac{d\rho}{ds} + \frac{1}{4}\frac{d}{ds}(\ln R) - \frac{a_R}{2a}\dot{\xi}^R\right)M^{-2}e^{2\rho}\dot{\xi}^R.$$

Now, recall from (7-7) that

$$-\frac{a_R}{2a} = \frac{1}{4R} + O(R^{-5/4}).$$

In particular, there exists some $R_2 > 0$ such that, for all s with $R(\xi(s)) > R_2$,

$$-\frac{a_R}{2a} \le \frac{1+\epsilon}{4R},$$

and we can assume that $R(\xi(s_0)) \ge R_2$. Thus, we have

$$\mu_R \le 2\left(\frac{d\rho}{ds} + \frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)\right)M^{-2}e^{2\rho}\dot{\xi}^R,$$

where the quantity in the parentheses is positive.

Thus, we conclude that

$$\mu_R^{1/2} \le \sqrt{2}M^{-1}\left(\frac{d\rho}{ds} + \frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)\right)^{\frac{1}{2}} e^{\rho}(\dot{\xi}^R)^{1/2} \le C\left(\frac{d\rho}{ds} + \frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)\right)e^{2\rho} + C\dot{\xi}^R.$$

It follows that

$$R^{-13/4-\epsilon}\mu_R^{1/2} \le CR^{-13/4-\epsilon}\left(\frac{d\rho}{ds} + \frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)\right)e^{2\rho} + CR^{-13/4-\epsilon}\dot{\xi}^R,$$

[3] We would like here to consider $d\mu/ds$, however, this would introduce the quantity a_θ, for which we do not directly have an evolution equation.

where the last term is clearly integrable since $\dot{\xi}^R = dR(\xi(s))/ds$ and $\frac{13}{4} - \epsilon > 1$. Finally, using (7-7) and an integration by parts to estimate the term containing $d\rho/ds$, we have, for any $s \in [s_0, s_1)$

$$\begin{aligned}\int_{s_0}^{s} R^{-13/4-\epsilon}\left(\frac{d\rho}{ds} + \frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)\right)e^{2\rho}\,ds \\ &= \int_{s_0}^{s} R^{-13/4-\epsilon}\frac{1}{2}\frac{de^{2\rho}}{ds}\,ds + \int_{s_0}^{s} R^{-13/4-\epsilon}\frac{1+\frac{\epsilon}{2}}{2}\frac{d}{ds}(\ln R)e^{2\rho}\,ds \\ &\le Ce^{2\rho}R^{-13/4-\epsilon} + C\int_{s_0}^{s} R^{-17/4-\epsilon}\dot{\xi}^R e^{2\rho}\,ds + C\int_{s_0}^{s} R^{-9/4}\dot{\xi}^R\,ds \le C.\end{aligned}$$

Thus, we have shown that $d(R^{-3/4-\epsilon}\dot{\xi}^R)/ds$ is integrable and, therefore, that $\dot{\xi}^R \le CR^{3/4+\epsilon}$ for some $C > 0$. This completes the proof of Theorem 8.1. □

9. Existence of initial data sets close to the asymptotic regime

In this section, we prove the following result:

Proposition 9.1 (existence of a class of initial data sets). *Fix $C_1 > 0$ and $\mathcal{A} \in [0, +\infty)$. For any $\epsilon > 0$, there exists $R_0 > 0$, $(U_0, U_1) \in H^1(S^1) \times L^2(S^1)$, $a_0 > 0 \in W^{2,1}(S^1)$ and $\eta_0 \in W^{1,1}(S^1)$ such that (U_0, U_1, a_0, η_0) satisfies the constraint equation* (2-7), *that is,*

$$\partial_\theta(\eta_0) = 2R_0U_1\,\partial_\theta(U_0), \tag{9-1}$$

and such that the conditions (6-1)–(6-5) *are all satisfied with $U(R_0, \theta) = U_0(\theta)$, $U_R(R_0, \theta) = U_1(\theta)$, $\eta(R_0, \theta) = \eta_0(\theta)$ and $a(R_0, \theta) = a_0(\theta)$. As a consequence, there exists an nonempty set of initial data satisfying* (6-1)–(6-5) *which is open in the natural topology associated with the initial data on $H^1(S^1) \times L^2(S^1) \times W^{2,1}(S^1) \times W^{1,1}(S^1)$.*

While our construction requires us to choose a sufficiently large R_0 (depending on ϵ), the ϵ satisfying the assumption of Theorem 7.1 depends only on a lower bound on R_0. Hence, the data constructed above satisfy the requirements of Theorem 7.1 provided R_0 is chosen sufficiently large.

Proof. Let $C_1 > 0$ and $\mathcal{A} \in [0, +\infty)$ be fixed. We define a_0 to be

$$a_0 = \frac{2\pi}{pR_0^{1/2}},$$

where $p > 0$ is a constant. Thus the associated term $\mathcal{P}$ reads $\mathcal{P} = pR_0^{1/2}$. We then define U_1 as

$$U_1 = \pm\frac{\mathcal{A}^{1/2}}{pR_0^{3/2}},$$

so that

$$\left(R_0\int_0^{2\pi} U_1 a_0^{-1}\,d\theta\right)^2 = \mathcal{A}.$$

Consider now any nonconstant $U_0 \in H^1(S^1)$. We will impose several conditions on U_0.

Let $\mathscr{E} = \int_{S^1} (a_0^{-1} U_1^2 + a_0 (U_0)_\theta^2)\, d\theta$ be the energy associated with our initial data set. Note that the energy correction[4] $\Gamma^U = (1/R_0) \int_{S^1} (U_0 - < U_0 >) U_1 a_0^{-1}\, dR$ equals 0 since $U_1 a_0^{-1}$ is constant.

Let $\mathscr{F} = \mathscr{P}\mathscr{E}$ and $\mathscr{G} = \mathscr{P}(\mathscr{E} + \Gamma^U)$ be the rescaled energy and the rescaled corrected energy associated with U_0, U_1 and a_0. Note that $\mathscr{G} = \mathscr{F}$ since $\Gamma^U = 0$, so that (5-25) trivially holds. Observe that

$$\mathscr{F} = \mathscr{P} \int_0^{2\pi} (a_0^{-1} U_1^2 + a_0 (U_0)_\theta^2)\, d\theta = \frac{\mathscr{A}}{R_0^2} + 2\pi \int_0^{2\pi} (U_0)_\theta^2\, d\theta.$$

Suppose now that $\int_0^{2\pi} (U_0)_\theta^2\, d\theta = C_1/(2\pi R_0)$, where $C_1 > 0$. Then we have

$$\mathscr{F} = \frac{\mathscr{A}}{R_0^2} + \frac{C_1}{R_0}.$$

In order to satisfy (6-2), we now fix p in terms of C_1 by setting

$$p = \left(\frac{2C_1}{5} \right)^{\frac{1}{2}}.$$

Then we compute

$$|c_1| = \left| \frac{2}{\sqrt{5}} - \frac{\mathscr{P}}{R_0 \mathscr{G}^{1/2}} \right| = \left| \frac{2}{\sqrt{5}} - \frac{\mathscr{P}}{R_0 \mathscr{F}^{1/2}} \right|.$$

On the other hand, we have

$$\frac{\mathscr{P}}{R_0 \mathscr{F}^{1/2}} = \frac{p R_0^{1/2}}{R_0 (\mathscr{A}/R_0^2 + C_1/R_0)^{1/2}} = \frac{2}{\sqrt{5}} \left(1 + \frac{\mathscr{A}}{R_0 C_1} \right)^{-1}.$$

This shows that (6-2) is satisfied provided that $\mathscr{A}/(R_0 C_1)$ is sufficiently small, which we can always ensure by choosing R_0 sufficiently large compared to $\mathscr{A}/C_1$.

One can then easily check that (6-1) and (6-5) hold true provided R_0 is sufficiently large. It remains to define η_0 so that (6-3) and the constraint equation (9-1) is satisfied.

For (6-3), we only need to ensure that $\left| \mathscr{Q}/R_0^3 \mathscr{F}^{1/2} - \frac{1}{\sqrt{5}} \right| \le \epsilon$. Recall that $\mathscr{Q} = \int_0^{2\pi} \frac{1}{2} K^2 e^{2\eta_0} a_0^{-1}\, d\theta$. By fixing $\eta_0(\theta = 0)$ we can certainly ensure that

$$\frac{K^2}{2} e^{2\eta_0(0)} a_0^{-1} = \frac{1}{2\pi\sqrt{5}} R_0^3 \mathscr{F}^{1/2}.$$

Now we define η_0 for all other values of θ, so that (9-1) is satisfied:

$$\eta_0(\theta) = \eta_0(0) + 2R_0 \int_0^\theta U_1 (U_0)_\theta\, d\theta = \eta_0(0) + 2R_0 U_1 (U_0(\theta) - U_0(0)).$$

From the above, we see that $\eta_0 \in W^{1,1}(S^1)$ (and in fact it is in $H^1(S^1)$) and that

$$|\eta_0(0) - \eta_0(\theta)| \le \int_0^{2\pi} |\eta_\theta|\, d\theta \le R_0 \frac{\mathscr{F}}{\mathscr{P}} \le \frac{1}{p R_0^{1/2}} R_0 \left(\frac{\mathscr{A}}{R_0^2} + \frac{C_1}{R_0} \right) \le \epsilon^2,$$

[4] We would like to thank an anonymous referee for pointing out this nice simplification.

again by choosing R_0 sufficiently large, depending only on C_1 and $\mathscr{A}$. We then check that

$$\left|\frac{\mathfrak{Q}}{R_0^3\mathscr{F}^{1/2}}-\frac{1}{\sqrt{5}}\right|=\frac{1}{R_0^3\mathscr{F}^{1/2}}\left|\mathfrak{Q}-2\pi\frac{K^2}{2}e^{2\eta_0(0)}a_0^{-1}\right|$$
$$\leq\frac{1}{R_0^3\mathscr{F}^{1/2}}\frac{K^2}{2}e^{2\eta_0(0)}a_0^{-1}\int_0^{2\pi}|e^{2(\eta_0(\theta)-\eta_0(0))}-1|\,d\theta\leq C\epsilon^2\leq\epsilon$$

provided ϵ is sufficiently small. □

Acknowledgments

The authors are very grateful to Hans Ringström for stimulating discussions on this subject and useful comments on a first version of this paper, and are also thankful to an anonymous referee for many constructive remarks. Part of this paper was written in the Fall Semester 2013 when LeFloch was a visiting professor at the Mathematical Sciences Research Institute (Berkeley) thanks to the support of the National Science Foundation under Grant No. 0932078 000. Part of this paper was written while Smulevici was a member of the Max Planck Institute for Gravitational Physics (Albert Einstein Institute). The authors were also partially supported by the Agence Nationale de la Recherche through the grant ANR SIMI-1-003-01.

References

[Berger et al. 1997] B. K. Berger, P. T. Chruściel, J. Isenberg, and V. Moncrief, “Global foliations of vacuum spacetimes with T^2 isometry”, *Ann. Physics* **260**:1 (1997), 117–148. MR 98j:83005 Zbl 0929.58013

[Choquet-Bruhat 2003] Y. Choquet-Bruhat, “Future complete S^1 symmetric Einsteinian spacetimes, the unpolarized case”, *C. R. Math. Acad. Sci. Paris* **337**:2 (2003), 129–136. MR 2004i:53104 Zbl 1027.83005

[Choquet-Bruhat and Moncrief 2001] Y. Choquet-Bruhat and V. Moncrief, “Future global in time Einsteinian spacetimes with U(1) isometry group”, *Ann. Henri Poincaré* **2**:6 (2001), 1007–1064. MR 2003g:53123 Zbl 0998.83007

[Christodoulou 1993] D. Christodoulou, “Bounded variation solutions of the spherically symmetric Einstein-scalar field equations”, *Comm. Pure Appl. Math.* **46**:8 (1993), 1131–1220. MR 95b:35221 Zbl 0853.35122

[Chruściel 1990] P. T. Chruściel, “On space-times with U(1) × U(1) symmetric compact Cauchy surfaces”, *Ann. Physics* **202**:1 (1990), 100–150. MR 91h:83007 Zbl 0727.53078

[Chruściel et al. 1990] P. T. Chruściel, J. Isenberg, and V. Moncrief, “Strong cosmic censorship in polarised Gowdy spacetimes”, *Classical Quantum Gravity* **7**:10 (1990), 1671–1680. MR 91i:83042 Zbl 0703.53081

[Grubic and LeFloch 2013] N. Grubic and P. G. LeFloch, “Weakly regular Einstein–Euler spacetimes with Gowdy symmetry: the global areal foliation”, *Arch. Ration. Mech. Anal.* **208**:2 (2013), 391–428. MR 3035982 Zbl 1266.83027

[Grubic and LeFloch 2015] N. Grubic and P. G. LeFloch, “On the area of the symmetry orbits in weakly regular Einstein–Euler spacetimes with Gowdy symmetry”, *SIAM J. Math. Anal.* **47**:1 (2015), 669–683. MR 3306012 Zbl 1320.83022

[LeFloch 2015] P. G. LeFloch, “An introduction to self-gravitating matter”, video lectures from graduate course, Institute Henri Poincaré, Paris, 2015, available at https://www.youtube.com/user/PoincareInstitute.

[LeFloch and Mardare 2007] P. G. LeFloch and C. Mardare, “Definition and stability of Lorentzian manifolds with distributional curvature”, *Port. Math.* (*N.S.*) **64**:4 (2007), 535–573. MR 2009e:53089 Zbl 1144.35311

[LeFloch and Rendall 2011] P. G. LeFloch and A. D. Rendall, “A global foliation of Einstein–Euler spacetimes with Gowdy-symmetry on T^3”, *Arch. Ration. Mech. Anal.* **201**:3 (2011), 841–870. MR 2012g:83008 Zbl 1256.35165

[LeFloch and Smulevici 2015] P. G. LeFloch and J. Smulevici, “Weakly regular T^2-symmetric spacetimes: the global geometry of future Cauchy developments”, *J. Eur. Math. Soc.* (*JEMS*) **17**:5 (2015), 1229–1292. MR 3346692

[LeFloch and Smulevici 2016] P. G. LeFloch and J. Smulevici, "Weakly regular T^2-symmetric spacetimes: the future causal geometry of Gowdy spacetimes", *J. Differential Equations* **260**:2 (2016), 1496–1521. MR 3419737

[LeFloch and Sormani 2015] P. G. LeFloch and C. Sormani, "The nonlinear stability of rotationally symmetric spaces with low regularity", *J. Funct. Anal.* **268**:7 (2015), 2005–2065. MR 3315585

[LeFloch and Stewart 2005] P. G. LeFloch and J. M. Stewart, "Shock waves and gravitational waves in matter spacetimes with Gowdy symmetry", *Port. Math.* (*N.S.*) **62**:3 (2005), 349–370. MR 2006g:83048

[LeFloch and Stewart 2011] P. G. LeFloch and J. M. Stewart, "The characteristic initial value problem for plane symmetric spacetimes with weak regularity", *Classical Quantum Gravity* **28**:14 (2011), 145019, 37. MR 2012j:83010 Zbl 1222.83028

[Ringström 2004] H. Ringström, "On a wave map equation arising in general relativity", *Comm. Pure Appl. Math.* **57**:5 (2004), 657–703. MR 2005h:58049 Zbl 1059.83011

[Ringström 2009] H. Ringström, "Strong cosmic censorship in T^3-Gowdy spacetimes", *Ann. of Math.* (2) **170**:3 (2009), 1181–1240. MR 2011d:83095 Zbl 1206.83115

[Ringström 2015] H. Ringström, "Instability of spatially homogeneous solutions in the class of $\mathbb{T}^2$-symmetric solutions to Einstein's vacuum equations", *Comm. Math. Phys.* **334**:3 (2015), 1299–1375. MR 3312437 Zbl 1310.83028

[Smulevici 2011] J. Smulevici, "On the area of the symmetry orbits of cosmological spacetimes with toroidal or hyperbolic symmetry", *Anal. PDE* **4**:2 (2011), 191–245. MR 2012k:83006 Zbl 1267.83013

Received 9 Aug 2014. Revised 14 Oct 2015. Accepted 6 Dec 2015.

PHILIPPE G. LEFLOCH: contact@philippelefloch.org
Laboratoire Jacques-Louis Lions and Centre National de la Recherche Scientifique, Université Pierre et Marie Curie (Paris 6), 4 Place Jussieu, 75252 Paris, France

JACQUES SMULEVICI: jacques.smulevici@math.u-psud.fr
Laboratoire de Mathématiques, Université Paris-Sud 11, Bât. 425, 91405 Orsay, France

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.397

OBSTACLE PROBLEM WITH A DEGENERATE FORCE TERM

KAREN YERESSIAN

We study the regularity of the free boundary at its intersection with the line $\{x_1 = 0\}$ in the obstacle problem

$$\triangle u = |x_1| \chi_{\{u>0\}} \quad \text{in } D,$$

where $D \subset \mathbb{R}^2$ is a bounded domain such that $D \cap \{x_1 = 0\} \neq \varnothing$.

We obtain a uniform $C^{1,1}$ bound on cubic blowups; we find all homogeneous global solutions; we prove the uniqueness of the blowup limit; we prove the convergence of the free boundary to the free boundary of the blowup limit; at the points with lowest Weiss balanced energy, we prove the convergence of the normal of the free boundary to the normal of the free boundary of the blowup limit and that locally the free boundary is a graph; and, finally, for a particular case we prove that the free boundary is not $C^{1,\alpha}$ regular near to a degenerate point for any $0 < \alpha < 1$.

1. Introduction

Let $D \subset \mathbb{R}^2$ be a bounded domain such that $D \cap \{x_1 = 0\} \neq \varnothing$. Let $g \in H^1(D)$ such that $g \geq 0$ on ∂D. Let $u \in H^1(D)$ be the unique minimiser of the functional

$$\int_D (|\nabla u|^2 + 2|x_1|u)\, dx \tag{1-1}$$

in the admissible set of functions

$$\{u \geq 0 \text{ a.e. in } D \text{ and } u = g \text{ on } \partial D\}.$$

For the existence and uniqueness of the minimiser u one may refer to [Kinderlehrer and Stampacchia 1980].

It is known (see [Petrosyan et al. 2012]) that $u \in C^{1,1}_{\text{loc}}(D)$ and

$$\triangle u = |x_1| \chi_{\{u>0\}} \quad \text{in } D \tag{1-2}$$

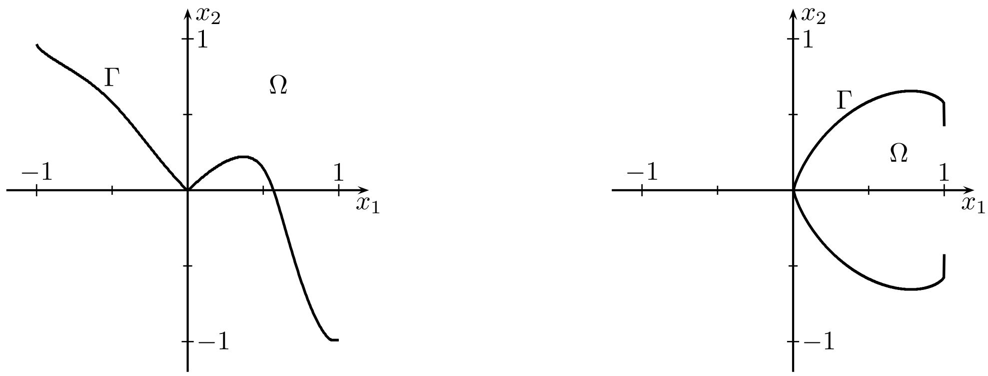
in the sense of distributions.

Let us denote by Ω the noncoincidence set and by Γ the free boundary, i.e.,

$$\Omega = \{x \in D \mid u(x) > 0\} \quad \text{and} \quad \Gamma = D \cap \partial\Omega.$$

Let us consider two examples. Set $D = (-1, 1)^2$. For the first example we take $g(x) = \frac{1}{16}(x_1 + x_2)^+$ and for the second example we take $g(x) = x_1^+(c - |x_2|)^+$, where $c \approx 0.42559$. The noncoincidence set and the free boundary are depicted in Figure 1 for both examples.

MSC2010: primary 35R35; secondary 35J60.
Keywords: free boundary, obstacle problem, degenerate, blowup, regularity.



(a) $g(x) = \frac{1}{16}(x_1 + x_2)^+$. (b) $g(x) = x_1^+(c - |x_2|)^+$ with $c \approx 0.42559$.

Figure 1. Ω and Γ in the examples.

In the case of the nondegenerate obstacle problem, i.e., when instead of $|x_1|$ we have f satisfying $f \geq c$ in D for some $c > 0$, the Lipschitz and C^1 regularity of the free boundary was proved for the first time in [Caffarelli 1977]. A good reference for nondegenerate obstacle problems is [Caffarelli 1998] and a good reference for obstacle-type problems is [Petrosyan et al. 2012].

In [Yeressian 2015], for a class of degenerate obstacle problems the optimal nondegeneracy of the solution is obtained. The proof of the optimal nondegeneracy is based on specially constructed comparison functions using harmonic polynomials. In this paper the nondegeneracy result in [Yeressian 2015] will be used numerous times.

Our approach to prove the regularity of the free boundaries is based on some directional monotonicity properties satisfied by the solutions. This method is based on the proof of C^1 regularity in [Petrosyan et al. 2012] and is closely related to [Alt 1977].

We use Hopf's lemma to prove the irregularity of the free boundary in a particular case which corresponds to the free boundary near to the origin in the example depicted in Figure 1(b). A related irregularity result has been proved in [Shahgholian et al. 2007], where the authors considered a two-phase membrane problem and in higher dimensions they proved that the free boundary is, in a neighbourhood of each branch point, the union of two C^1-graphs, but in general these graphs are not $C^{1,\text{Dini}}$ ($C^{1,\text{Dini}}$ includes all $C^{1,\alpha}$ for $0 < \alpha < 1$).

Studying obstacle problems with a degenerate force term reveals rather unexpected behaviour of the solution, such as the fact that the free boundary usually forms a certain angle at its intersections with the line $\{x_1 = 0\}$, where the force term is degenerate.

In the problem of the free boundary near contact points with the fixed boundary — see [Shahgholian and Uraltseva 2003] — where the solution satisfies a homogeneous Dirichlet boundary condition, a similar strong influence of the data of the problem on the structure of the free boundary has been observed.

Varvaruca and Weiss [2011; 2012; 2014] have studied 2-dimensional or axisymmetric, 3-dimensional, inviscid, incompressible fluids acted on by gravity and with a free surface. These problems are in the class of Bernoulli free boundary problems, but the degeneracies in the force terms give rise to similar situations as encountered in this paper and has been a motivation for considering the problem in this paper.

This paper is structured as follows. In Section 2, the main results of this paper are presented. In Section 3, we prove uniform $C^{1,1}$ bounds on cubic blowups. In Section 4, using the Weiss balanced energy we prove the homogeneity of the blowup limits. In Section 5, we classify all possible homogeneous global solutions. In Section 6, using the fact that possible blowup limits form a discrete set we prove the uniqueness of the blowup limits. In Section 7, using a lower bound for homogeneous global solutions and the optimal nondegeneracy result in [Yeressian 2015] we prove the convergence of the free boundary to the free boundary of the blowup limit. In Section 8, we prove the convergence of the normal of the free boundary to the normal of the free boundary of the blowup limit at regular points. In Section 9, we prove that in a neighbourhood of a regular point the free boundary can be given as a graph. In Section 10, we prove that under some assumptions the free boundary near to a degenerate point is either flat or not $C^{1,\alpha}$ for any $0 < \alpha < 1$. In Section 11, we conclude this paper with a discussion about further directions of research on obstacle problems with degenerate forces.

2. Main results

Let us define a cubic blowup of u as follows:

Definition 1. Let $B_{r_0} \subset D$, then we define, for $0 < r < r_0$,

$$u_r(x) = \frac{u(rx)}{r^3} \quad \text{for } x \in B_1$$

and call u_r the (cubic) blowup of u at 0.

In the following theorem we prove that for $r > 0$ the family u_r is uniformly bounded in $C^{1,1}(B_1)$.

Theorem 2 (uniform $C^{1,1}$ bounds on blowups). *There exists a $C > 0$ such that, if u is a solution in D, $r_0 > 0$, $B_{r_0} \subset D$ and $0 \in \Gamma$, then we have the estimate*

$$\|u_r\|_{C^{1,1}(B_1)} \leq C \tag{2-1}$$

for $0 < r < \frac{1}{6} r_0$.

The proof of this theorem is based on the optimal growth result proved in [Yeressian 2015].

From the uniform bound (2-1) it follows that, for any sequence r_j such that $r_j \to 0$, there exists a subsequence r_{j_k} and $v \in C^{1,1}(B_1)$ such that $u_{j_k} \to v$ in $C^1(B_1)$.

Let us consider for $u \in H^1(B_r)$ the Weiss balanced energy

$$W(r, u) = \frac{1}{r^6} \int_{B_r} (|\nabla u|^2 + 2|x_1|u)\, dx - \frac{3}{r^7} \int_{\partial B_r} u^2 \, d\sigma(x). \tag{2-2}$$

The Weiss balanced energy [1998; 1999] was introduced to study the free boundary in the nondegenerate obstacle problem. The energy in (2-2) has been adapted to the first-order homogeneity of the force term $|x_1|$. For the Weiss balanced energy for different homogeneities, one may refer to [Petrosyan et al. 2012].

As we will see, for u a solution in D with $0 \in D$, by a monotonicity result for the Weiss balanced energy, the right limit $W(+0, u)$ exists but might be $-\infty$. If $0 \in \Gamma$ then $W(+0, u) > -\infty$.

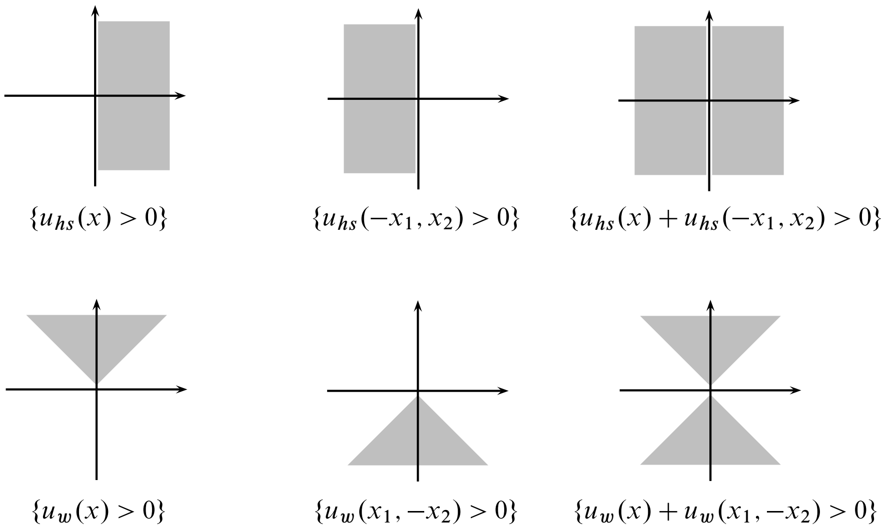


Figure 2. The only possible noncoincidence sets of nontrivial homogeneous global solutions.

Definition 3. Let u be a solution in D, $0 \in D$ and $0 \in \Gamma$. Then we call $v \in C^{1,1}(B_1)$ a blowup limit if there exists $r_j \to 0$ such that $u_{r_j} \to v$ in $C^1(B_1)$.

Using the Weiss balanced energy, if v is a blowup limit at 0 then v is a third-order homogeneous global solution and $W(+0, u) = W(1, v)$.

So we are led to find all the solutions of the obstacle problem

$$\begin{cases} \triangle u = |x_1| \chi_{\{u>0\}} \text{ in } \mathbb{R}^2, \\ u \text{ third-order homogeneous.} \end{cases} \tag{2-3}$$

Clearly $u = 0$ is a trivial solution of (2-3).

Let us define

$$u_{hs}(x) = \tfrac{1}{6}(x_1^+)^3 \quad \text{and} \quad u_w(x) = \left(\tfrac{1}{6}|x_1|^3 + \tfrac{1}{12}x_2^3 - \tfrac{1}{4}x_1^2 x_2\right)\chi_{\{x_2>|x_1|\}}. \tag{2-4}$$

Theorem 4 (classification of homogeneous global solutions). *The only nontrivial solutions of* (2-3) *are* u_w, $u_w(x_1, -x_2)$, $u_w + u_w(x_1, -x_2)$, u_{hs}, $u_{hs}(-x_1, x_2)$ *and* $u_{hs} + u_{hs}(-x_1, x_2)$.

To prove Theorem 4 we first find all the solutions of the corresponding no-sign obstacle problem and then among these solutions we find the nonnegative ones.

All possible noncoincidence sets of nontrivial homogeneous global solutions, i.e., the noncoincidence sets of the nontrivial solutions of (2-3), are depicted in Figure 2.

It is easy to see that $W(1, u_w) = W(1, u_w(x_1, -x_2))$, $W(1, u_w + u_w(x_1, -x_2)) = 2W(1, u_w)$, $W(1, u_{hs}) = W(1, u_{hs}(-x_1, x_2))$, $W(1, u_{hs} + u_{hs}(-x_1, x_2)) = 2W(1, u_{hs})$ and, by direct computation, we see that $0 < W(1, u_w)$ and

$$2W(1, u_w) < W(1, u_{hs}).$$

So we have the following four possible energy levels together with the order between them:

$$W(1,u_w) < 2W(1,u_w) < W(1,u_{hs}) < 2W(1,u_{hs}).$$

Let us define, for $y \in \Gamma \cap \{x_1 = 0\}$ and $r > 0$,

$$W(r, y, u) = W(r, u(\cdot + y)). \tag{2-5}$$

Based on the four possible values of $W(+0, x, u)$ (the value 0 is excluded by the nondegeneracy) for $x \in \Gamma \cap \{x_1 = 0\}$, the points of $\Gamma \cap \{x_1 = 0\}$ get classified in four types.

Definition 5. We call $y \in \Gamma \cap \{x_1 = 0\}$ a degenerate free boundary point if there exists $r_j \to 0$ such that $u(\cdot + y)_{r_j} \to u_{hs}$ or $u(\cdot + y)_{r_j}(x) \to u_{hs}(-x_1, x_2)$ in $C^1(B_1)$.

We use this name for points where a blowup limit is u_{hs} or $u_{hs}(-x_1, x_2)$ by following the naming for similar points in the problem studied in [Varvaruca and Weiss 2011].

In the example depicted in Figure 1(b), the origin is a degenerate free boundary point with u_{hs} as a blowup limit.

By our uniform bounds on the blowups it follows that 0 is degenerate if and only if $W(+0, u) = W(1, u_{hs})$.

Definition 6. We call $y \in \Gamma \cap \{x_1 = 0\}$ a regular free boundary point if there exists $r_j \to 0$ such that $u(\cdot + y)_{r_j} \to u_w$ or $u(\cdot + y)_{r_j}(x) \to u_w(x_1, -x_2)$ in $C^1(B_1)$.

In the example depicted in Figure 1(a) a point close to the origin is a regular free boundary point with u_w as a blowup limit.

By our uniform bounds on the blowups it follows that 0 is regular if and only if $W(+0, u) = W(1, u_w)$, i.e., it has the lowest Weiss balanced energy.

Theorem 7 (uniqueness of blowup limits). *If u is a solution in D, $0 \in D$ and $0 \in \Gamma$ then the blowup limit at the origin is unique.*

Let us define, for $\delta > 0$ and $k = 0, 1$,

$$\sigma_k(\delta) = \sup_{0<r\le\delta} \|u_r - u_0\|_{C^k(B_1)}, \tag{2-6}$$

where u_0 is the unique blowup limit.

Theorem 8 (convergence of the free boundary). *There exists $C_1 > 0$ and $C_2 > 0$ such that if u is a solution in D, $0 \in D$ and $0 \in \Gamma$ then, for $x \in \Gamma$ close enough to the origin, if $W(+0, u) \in \{W(1, u_w), 2W(1, u_w)\}$ then we have*

$$d(x, \Gamma_{u_0}) \le C_1 \big(\sigma_0(C_2|x|)\big)^{1/2} |x|, \tag{2-7}$$

where Γ_{u_0} is the free boundary of the unique blowup limit, and, if $W(+0, u) \in \{W(1, u_{hs}), 2W(1, u_{hs})\}$, then

$$|x_1| \le C_1 \big(\sigma_0(C_2|x|)\big)^{1/3} |x|. \tag{2-8}$$

The proof of this theorem is based on a lower bound for the nontrivial homogeneous global solutions and the nondegeneracy result proved in [Yeressian 2015].

From Theorem 8, it follows that all points of $\Gamma\cap\{x_1=0\}\cap\{W(+0,x,u)\in\{W(1,u_w),2W(1,u_w)\}\}$ are isolated points of $\Gamma\cap\{x_1=0\}$ (in the topology of $\{x_1=0\}$), in particular.

Theorem 9 (convergence of normals and the free boundary as a graph at regular points). *There exists $C_1>0$ and $C_2>0$ such that if u is a solution in D, $0\in D$ and $0\in\Gamma$ is a regular free boundary point with blowup limit u_w then there exists $\epsilon>0$ and*

$$\gamma\in C\big(-\tfrac14\epsilon,\tfrac14\epsilon\big)\cap C^1\big(\big(-\tfrac14\epsilon,\tfrac14\epsilon\big)\backslash\{0\}\big)$$

such that

$$\begin{aligned}\Gamma\cap\big\{|x_1|<\tfrac14\epsilon\big\}\cap B_\epsilon=\big\{(x_1,\gamma(x_1))\;\big|\;x_1\in\big(-\tfrac14\epsilon,\tfrac14\epsilon\big)\big\},&\\ \big|\gamma(x_1)-|x_1|\big|\le C_1\big(\sigma_0(C_2|x_1|)\big)^{1/2}|x_1| \qquad &\text{for } x_1\in\big(-\tfrac14\epsilon,\tfrac14\epsilon\big),\\ \left|\gamma'(x_1)-\frac{x_1}{|x_1|}\right|\le C_1\big(\sigma_1(C_2|x_1|)\big)^{1/2} \qquad &\text{for } x_1\in\big(-\tfrac14\epsilon,\tfrac14\epsilon\big)\backslash\{0\}.\end{aligned}$$

The proof of this theorem is mainly based on a directional monotonicity result proved in Lemma 37. There we prove that $\partial_\nu u\ge 0$ in $B_r(x)$ for $x\in\Gamma\cap\{x_1>0\}\cap\partial B_{1/4}$ if, for a given $\nu\in\partial B_1$ with $\nu\cdot\nu_w>0$, r is small enough and u is close enough to u_w in $C^1(B_1)$. The vector ν_w is the normal to the free boundary of u_w in the half-plane $\{x_1>0\}$, pointing into the noncoincidence set of u_w. This directional monotonicity result establishes the convergence of the normal of the free boundary to the normal of the free boundary of the blowup limit.

As we will see, from Theorem 9 it follows that, in the case when the origin is a regular point but with $u_w(x_1,-x_2)$ as blowup limit, and in the case when $W(+0,u)=2W(1,u_w)$, the free boundary is a graph or the union of two graphs, respectively.

In the following theorem, in particular cases we show that the free boundary near to a degenerate point is not $C^{1,\alpha}$ smooth.

Theorem 10 (an irregularity result at degenerate points). *Let u be a solution in D with $0\in D$. Suppose also that there exists $\delta>0$ such that $B_\delta\subset D$, $\partial_{x_2}u\le 0$ in $B_\delta\cap\{x_1>0,x_2>0\}$, there exists $\rho\in C\big(\big[0,\tfrac12\delta\big)\big)\cap C^1\big(\big[0,\tfrac12\delta\big)\big)$ such that $\rho(0)=\rho'(+0)=0$, $\rho\ge 0$ in $\big(0,\tfrac12\delta\big)$, ρ is convex and*

$$\Omega\cap B_\delta\cap\big\{x_1>0,\;0<x_2<\tfrac12\delta\big\}=B_\delta\cap\big\{0<x_2<\tfrac12\delta,\;\rho(x_2)<x_1\big\}.$$

Then either $\rho=0$ and $u=u_{hs}$ in $\Omega\cap B_\delta\cap\big\{x_1>0,\;0<x_2<\tfrac12\delta\big\}$ or the free boundary part $\Gamma\cap\{x_1>0\}$ is not $C^{1,\alpha}$ regular at 0 for any $0<\alpha<1$.

Let us note that the conditions in this theorem correspond to the example depicted in Figure 1(b).

The proof of this theorem relies on considering the nonnegative function $v=-\partial_{x_2}u$ and using the quantitative Hopf lemma (see [Han and Lin 2011]).

3. Uniform bounds on blowups

The following theorem is a consequence of the Harnack inequality and is a special case of the optimal growth theorem in [Yeressian 2015].

Theorem 11. *There exists a $C > 0$ such that if $B_r(y) \subset D$ then we have*

$$u(x) \le C(u(y) + r^2(r + |y_1|)) \quad \textit{for } x \in B_{r/2}(y).$$

Based on this optimal growth estimate, in the following theorem we prove an estimate on the growth of the solution near the free boundary.

Lemma 12. *There exists a $C > 0$ such that if u is a solution in D, $y \in \Omega$, $d = d(y, \Gamma)$ and $B_{5d}(y) \subset D$ then*

$$u(x) \le Cd^2(d + |y_1|) \quad \textit{for } x \in B_d(y). \tag{3-1}$$

Proof. Let $z \in \Gamma$ be such that $d = |y - z|$. We have, for $r = 4d$,

$$B_r(z) = B_{4d}(z) \subset B_{4d+|y-z|}(y) = B_{5d}(y) \subset D.$$

By Theorem 11 we have that, because $z \in \Gamma$ and $B_r(z) \subset D$,

$$u(x) \le C_1 r^2(r + |z_1|) \quad \text{for } x \in B_{r/2}(z). \tag{3-2}$$

We have

$$B_d(y) \subset B_{d+|y-z|}(z) = B_{2d}(z) = B_{r/2}(z). \tag{3-3}$$

By (3-2) and (3-3) we obtain

$$\begin{aligned} u(x) \le C_1 r^2(r + |z_1|) = C_1(4d)^2(4d + |z_1|) &\le C_2 d^2(d + |z_1|) \\ &\le C_2 d^2(d + |z_1 - y_1| + |y_1|) \\ &\le C_2 d^2(2d + |y_1|) \le C_3 d^2(d + |y_1|) \quad \text{for } x \in B_d(y), \end{aligned}$$

which proves the lemma. □

Let us define

$$\psi(t) = \tfrac{1}{6}|t|^3 \quad \text{for } t \in \mathbb{R} \tag{3-4}$$

and, for $t_0 \in \mathbb{R}$,

$$w_{t_0}(t) = \psi(t) - \psi(t_0) - \psi'(t_0)(t - t_0) \quad \text{for } t \in \mathbb{R}.$$

Then there exists $C > 0$ such that for $t, t_0 \in \mathbb{R}$ we have

$$w_{t_0}(t) \le C|t - t_0|^2(|t_0| + |t - t_0|). \tag{3-5}$$

Proof of Theorem 2. We have

$$\|u_r\|_{L^\infty(B_1)} = \frac{1}{r^3}\|u\|_{L^\infty(B_r)}, \quad \|\nabla u_r\|_{L^\infty(B_1)} = \frac{1}{r^2}\|\nabla u\|_{L^\infty(B_r)}, \quad [\nabla u_r]_{C^{0,1}(B_1)} = \frac{1}{r}[\nabla u]_{C^{0,1}(B_r)}.$$

So, if we prove that for some $C > 0$ we have

$$\|u\|_{L^\infty(B_r)} \le Cr^3, \tag{3-6}$$

$$\|\nabla u\|_{L^\infty(B_r)} \le Cr^2, \tag{3-7}$$

$$[\nabla u]_{C^{0,1}(B_r)} \le Cr, \tag{3-8}$$

then the lemma is proved.

There exists $C > 0$ such that for v a harmonic function in B_1 we have

$$|\nabla v(0)| \le C\|v\|_{L^\infty(B_1)} \quad \text{and} \quad [\nabla v]_{C^{0,1}(B_{1/2})} \le C\|v\|_{L^\infty(B_1)}.$$

By scaling we obtain that for v harmonic in B_η we have

$$|\nabla v(0)| \le \frac{C}{\eta}\|v\|_{L^\infty(B_\eta)} \tag{3-9}$$

$$[\nabla v]_{C^{0,1}(B_{\eta/2})} \le \frac{C}{\eta^2}\|v\|_{L^\infty(B_\eta)}. \tag{3-10}$$

For $x \in \Omega$ let $d = d(x, \Gamma)$; then we have

$$B_{5d}(x) \subset B_{5d+|x|} \subset B_{5|x|+|x|} = B_{6|x|},$$

so if $x \in B_{(1/6)r_0}$ then $B_{5d}(x) \subset D$.

Now, by Lemma 12, we obtain that for $x \in B_{(1/6)r_0}$ we have

$$\|u\|_{L^\infty(B_d(x))} \le Cd^2(d + |x_1|). \tag{3-11}$$

Let $0 < r < \frac{1}{6}r_0$.

To prove (3-6), we compute, for $x \in B_r$,

$$|u(x)| \le \|u\|_{L^\infty(B_d(x))} \le Cd^2(d + |x_1|) \le C|x|^2(|x| + |x_1|) = 2C|x|^3 \le 2Cr^3.$$

To prove (3-7), using $w'_{x_1}(x_1) = 0$, (3-9), (3-11) and (3-5), we compute, for $x \in B_r$,

$$\begin{aligned} |\nabla u(x)| = |\nabla(u - w_{x_1})(x)| &\le \frac{C_1}{d}\|u - w_{x_1}\|_{L^\infty(B_d(x))} \\ &\le \frac{C_1}{d}\|u\|_{L^\infty(B_d(x))} + \frac{C_1}{d}\|w_{x_1}\|_{L^\infty(B_d(x))} \\ &\le C_2 d(d + |x_1|) + C_3 d(d + |x_1|) = C_4 d(d + |x_1|). \end{aligned} \tag{3-12}$$

From (3-12) it follows that

$$|\nabla u(x)| \le 2C_4|x|^2 \le 2C_4 r^2. \tag{3-13}$$

It remains to prove (3-8). We should show that

$$|\nabla u(x) - \nabla u(y)| \le Cr|x - y| \quad \text{for all } x, y \in B_r.$$

Fix $x, y \in B_r$. In the case $B_{|x-y|}\big(\frac{1}{2}(x + y)\big) \subset \Omega$ let us denote $z = \frac{1}{2}(x + y)$. We have $d = d(z, \Gamma) \ge |x - y|$.

By (3-10) and (3-11), we compute

$$\begin{aligned}\frac{|\nabla u(x)-\nabla u(y)|}{|x-y|} &\le [\nabla u]_{C^{0,1}(B_{|x-y|/2}(z))}\\ &\le [\nabla u]_{C^{0,1}(B_{d/2}(z))}\\ &\le [\nabla(u-w_{z_1})]_{C^{0,1}(B_{d/2}(z))}+[w'_{z_1}]_{C^{0,1}(B_{d/2}(z))}\\ &\le \frac{C_1}{d^2}\|u-w_{z_1}\|_{L^\infty(B_d(z))}+[w_{z_1}]_{C^2(B_{d/2}(z))}\\ &\le \frac{C_1}{d^2}\|u\|_{L^\infty(B_d(z))}+\frac{C_1}{d^2}\|w_{z_1}\|_{L^\infty(B_d(z))}+[\psi]_{C^2(B_{d/2}(z))}\\ &\le \frac{C_1}{d^2}C_2d^2(d+|z_1|)+\frac{C_1}{d^2}C_3d^2(d+|z_1|)+C_4(d+|z_1|)\\ &= C_5(d+|z_1|)\\ &\le 2C_5r.\end{aligned}$$

In the case $B_{|x-y|}\big(\frac{1}{2}(x+y)\big)\cap\Omega^c\neq\varnothing$, by (3-12) we compute

$$\begin{aligned}|\nabla u(x)-\nabla u(y)| &\le |\nabla u(x)|+|\nabla u(y)|\\ &\le Cd(x,\Gamma)(d(x,\Gamma)+|x_1|)+Cd(y,\Gamma)(d(y,\Gamma)+|y_1|)\\ &\le \tfrac{3}{2}C|x-y|(d(x,\Gamma)+|x_1|)+\tfrac{3}{2}C|x-y|(d(y,\Gamma)+|y_1|)\\ &\le C_1r|x-y|\end{aligned}$$

and this finishes the proof of the theorem. □

4. Homogeneity of blowup limits

Most of the results in this section are well known; one may refer to [Petrosyan et al. 2012; Weiss 1998; 1999]. But for the sake of completeness we include the proofs.

The Weiss balanced energy $W(r,u)$ is defined in (2-2).

Lemma 13. *For $r, s>0$ and $u\in H^1(B_{rs})$, we have $W(rs,u)=W(s,u_r)$.*

For $u\in H^1(B_{r_0})$, $W(r,u)$ as a function of $0<r<r_0$ is locally bounded and absolutely continuous.

For u a solution in B_{r_0} and $0<r<r_0$, we have

$$\frac{d}{dr}W(r,u)=2r\int_{\partial B_1}(\partial_r u_r)^2\,d\sigma(x). \tag{4-1}$$

For u a third-order homogeneous solution in B_1, we have

$$W(1,u)=\int_{B_1}|x_1|u\,dx. \tag{4-2}$$

Proof. Let $r, s > 0$ and $u \in H^1(B_{rs})$. We compute

$$\begin{aligned} W(rs,u) &= \frac{1}{(rs)^6}\int_{B_{rs}} (|\nabla u|^2 + 2|x_1|u)\,dx - \frac{3}{(rs)^7}\int_{\partial B_{rs}} u^2\,d\sigma(x) \\ &= \frac{1}{s^6}\frac{1}{r^4}\int_{B_s} \big(|\nabla u(rx)|^2 + 2r|x_1|u(rx)\big)\,dx - \frac{3}{s^7}\frac{1}{r^6}\int_{\partial B_s} u^2(rx)\,d\sigma(x) \\ &= \frac{1}{s^6}\int_{B_s} \big(|\nabla u_r(x)|^2 + 2|x_1|u_r\big)\,dx - \frac{3}{s^7}\int_{\partial B_s} u_r^2\,d\sigma(x) = W(s,u_r), \end{aligned}$$

which proves the first claim.

Let $u \in H^1(B_{r_0})$; then, for $0 < r < r_0$, by direct computation using polar coordinates we have

$$\int_{\partial B_r} u^2\,d\sigma(x) = -2r\int_{B_{r_0}\setminus B_r} \frac{1}{|x|^2}u(x)\nabla u(x)\cdot x\,dx + \frac{r}{r_0}\int_{\partial B_{r_0}} u^2(x)\,d\sigma(x). \tag{4-3}$$

The equation (4-3) together with the fact that if $f \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ then $\int_{B_r} f\,dx$ as a function of r is bounded and absolutely continuous proves the second claim.

Let u be a solution in B_{r_0}, then we have (see [Petrosyan et al. 2012]) $u \in C^{1,1}_{\mathrm{loc}}(B_{r_0})$. Let $0 < r < r_0$, then we compute

$$\begin{aligned} &\frac{1}{2}\frac{d}{dr}W(r,u) \\ &\quad= \frac{1}{2}\frac{d}{dr}W(1,u_r) \\ &\quad= \frac{1}{2}\bigg(\int_{B_1} \big(2\nabla u_r(x)\cdot\nabla\partial_r u_r(x) + 2|x_1|\partial_r u_r\big)\,dx - 6\int_{\partial B_1} u_r\partial_r u_r\,d\sigma(x)\bigg) \\ &\quad= \int_{B_1} \big(\nabla u_r(x)\cdot\nabla\partial_r u_r(x) + |x_1|\partial_r u_r\big)\,dx - 3\int_{\partial B_1} u_r\partial_r u_r\,d\sigma(x) \\ &\quad= \int_{B_1} \big(-\triangle u_r(x)\partial_r u_r(x) + |x_1|\partial_r u_r\big)\,dx + \int_{\partial B_1} \partial_\nu u_r(x)\partial_r u_r(x)\,d\sigma(x) - 3\int_{\partial B_1} u_r\partial_r u_r\,d\sigma(x) \\ &\quad= \int_{\partial B_1} \big(\partial_\nu u_r(x) - 3u_r\big)\partial_r u_r\,d\sigma(x). \end{aligned}$$

It is easy to see that on ∂B_1 we have

$$\partial_\nu u_r(x) - 3u_r = r\,\partial_r u_r,$$

which proves the third claim.

Let u be a solution in B_1. We compute

$$\begin{aligned} W(1,u) &= \int_{B_1} \big(|\nabla u(x)|^2 + 2|x_1|u\big)\,dx - 3\int_{\partial B_1} u^2\,d\sigma(x) \\ &= \int_{B_1} (-\triangle u(x))u(x)\,dx + \int_{\partial B_1} \partial_\nu u(x)u(x)\,d\sigma(x) + \int_{B_1} 2|x_1|u\,dx - 3\int_{\partial B_1} u^2\,d\sigma(x) \\ &= \int_{\partial B_1} \partial_\nu u(x)u(x)\,d\sigma(x) + \int_{B_1} |x_1|u\,dx - 3\int_{\partial B_1} u^2\,d\sigma(x) \end{aligned}$$

$$= \int_{B_1} |x_1| u \, dx + \int_{\partial B_1} (\partial_\nu u - 3u) u \, d\sigma(x).$$

For a third-order homogeneous function we have $\partial_\nu u = 3u$; thus the last integral is null and this proves the last claim. □

If u is a solution in B_{r_0} for some $r_0 > 0$ then, by (4-1), $W(r, u)$ is nondecreasing in $0 < r < r_0$; thus the limit $\lim_{r \to 0, r > 0} W(r, u) = W(+0, u)$ exists but might be $-\infty$. If $0 \in \Gamma$ then by Theorem 2 we have $\|u_r\|_{L^\infty(B_1)} \le C$ for small enough $0 < r$ and from this it follows that

$$-\frac{1}{r^7} \int_{\partial B_r} u^2 \, d\sigma(x) = -\int_{\partial B_1} u_r^2 \, d\sigma(x) \ge -c_1;$$

thus $W(r, u) \ge -3c_1$ and $W(+0, u) \ge -3c_1 > -\infty$.

For $y \in \Gamma \cap \{x_1 = 0\}$ and $r > 0$, $W(r, y, u)$ is defined in (2-5).

Lemma 14. *$W(+0, x, u)$ is an upper-semicontinuous function of $x \in \Gamma \cap \{x_1 = 0\}$.*

Proof. For $x \in \Gamma \cap \{x_1 = 0\}$, by the monotonicity of $W(r, x, u)$ as a function of $r > 0$ and its continuity as a function of x it follows that $W(+0, x, u) = \lim_{r \to 0, r > 0} W(r, x, u)$ is upper-semicontinuous in $\Gamma \cap \{x_1 = 0\}$. □

Assume v is a third-order homogeneous function in B_1, i.e., $v(0) = 0$ and $v(x) = v(x/(2|x|))(2|x|)^3$ for all $x \in B_1 \setminus \{0\}$. Then we might extend v as a third-order homogeneous function in $\mathbb{R}^2$ as $v(x) = v(x/(2|x|))(2|x|)^3$ for all $x \in B_1^c$. Let us note that the term on the right-hand side is well defined because for $x \in B_1^c$ we have $x/(2|x|) \in B_1$. From this definition of extension it follows that $v(rx) = r^3 v(x)$ for all $x \in \mathbb{R}^2$ and $r \ge 0$.

The following theorem is a special case of the main theorem in [Yeressian 2015].

Theorem 15. *There exists a $c > 0$ such that if u is a solution in D, $y \in \Omega$ and $B_r(y) \Subset D$ then we have*

$$\sup_{\Omega \cap \partial B_r(y)} u \ge u(y) + c r^2 (r + |y_1|).$$

A blowup limit is defined in Definition 3.

Lemma 16. *Let v be a blowup limit. Then v is a third-order nontrivial homogeneous solution in B_1, the third-order homogeneous extension of v in $\mathbb{R}^2$ is a global solution, and $W(+0, u) = W(r, v)$ for $r > 0$.*

Proof. Assume $v \in C^{1,1}(B_1)$ is a blowup limit and $u_{r_j} \to v$ in $C^1(B_1)$.

From $u_{r_j} \ge 0$ in B_1 it follows that $v \ge 0$ in B_1. By the convergence $u_{r_j} \to v$ in $C^1(B_1)$ it follows that $\triangle u_{r_j} \to \triangle v$ in $H^{-1}(B_1)$ and in particular as distributions. Also $\chi_{\{u_{r_j} > 0\}} \to \chi_{\{v > 0\}}$ in $L^1(B_1)$ and thus $|x_1| \chi_{\{u_{r_j} > 0\}} \to |x_1| \chi_{\{v > 0\}}$ as distributions. Now (1-2) holds for u_{r_j} in B_1, so passing to the limit as $j \to \infty$ we obtain that v satisfies (1-2) in B_1. This together with $v \ge 0$ in B_1 proves that v is a solution to the obstacle problem in B_1.

For $0 < s < 1$ we compute

$$W(+0, u) = \lim_{j \to \infty} W(s r_j, u) = \lim_{j \to \infty} W(s, u_{r_j}) = W(s, v). \tag{4-4}$$

Thus $W(s, v)$ is independent of $0 < s < 1$.

Now, by (4-1), we obtain that for $0 < s < 1$

$$0 = \frac{d}{ds} W(s, v) = 2s \int_{\partial B_1} (\partial_s v_s)^2 \, d\sigma(x).$$

From here it follows that $\nabla v \cdot x - 3v = 0$ in B_1 and hence v is third-order homogeneous in B_1.

Now let us prove that v is not 0 in B_1, i.e., v is nontrivial.

Let $\delta > 0$ and $B_\delta \subset D$. Let $0 < r < \delta$ and $y \in B_{r/2} \cap \Omega$; then we have

$$B_{r/4}(y) \subset B_{r/4+|y|} \subset B_{r/4+r/2} = B_{3r/4} \Subset D,$$

thus by Theorem 15 we have

$$\sup_{\partial B_{r/4}(y)} u \geq u(y) + c\left(\tfrac{1}{4} r\right)^3.$$

We compute

$$\partial B_{r/4}(y) \subset B_{r/2}(y) \subset B_{r/2+|y|} \subset B_r,$$

so we have

$$\sup_{B_r} u \geq \sup_{\partial B_{r/4}(y)} u \geq u(y) + c\left(\tfrac{1}{4} r\right)^3 \geq \tfrac{1}{4^3} c r^3$$

and thus

$$\sup_{B_1} u_r \geq \tfrac{1}{4^3} c.$$

From this inequality, taking $r = r_j \to 0$, we obtain that v is not identically 0 in B_1.

Let us again denote by v the extension of v in $\mathbb{R}^2$. Then it is easy to see that, because v is a solution in B_1 and $v(rx) = r^3 v(x)$ for $x \in \mathbb{R}^2$ and $r \geq 0$, v is a solution in $\mathbb{R}^2$, i.e., a global solution.

By third-order homogeneity of v we have $W(r, v) = W\left(\tfrac{1}{2}, v\right)$ for $r > 0$ and this together with (4-4) proves the last claim of the lemma. □

5. Homogeneous global solutions

In this section we classify all possible solutions of the problem (2-3). The solutions of (2-3) form the subset of nonnegative solutions of the following no-sign obstacle problem (see [Petrosyan et al. 2012] for more on no-sign obstacle problems)

$$\begin{cases} \triangle u = |x_1| \chi_{\Omega(u)} & \text{in } \mathbb{R}^2, \\ \Omega(u) = \{u = |\nabla u| = 0\}^c, \\ u \ \text{third-order homogeneous.} \end{cases} \tag{5-1}$$

We first classify the nontrivial solutions of (5-1) and then find the subset of nonnegative and nontrivial solutions of (5-1), and thus obtain the classification of the nontrivial solutions of the problem (2-3).

In the rest of this section we always assume that $u \neq 0$ in $\mathbb{R}^2$, i.e., we discuss only the nontrivial solutions, so $\Omega \neq \varnothing$.

In both problems, by homogeneity, the set Ω is an open cone in $\mathbb{R}^2\backslash\{0\}$, i.e., for $x \in \Omega$ and $r > 0$ we have $rx \in \Omega$.

Either Ω is equal to $\mathbb{R}^2\backslash\{0\}$ or it is at most a countable union of disjoint connected open cones in $\mathbb{R}^2\backslash\{0\}$.

To classify the solutions in both problems we first establish if there exists a solution with $\Omega = \mathbb{R}^2\backslash\{0\}$. Then we find all the connected cones Ω not equal to $\mathbb{R}^2\backslash\{0\}$ for which there exists a corresponding solution.

Let us define

$$U(\theta) = u(e^{i\theta}) - \tfrac{1}{3} i\, \partial_\theta u(e^{i\theta}).$$

Lemma 17. *If u is a third-order homogeneous function in a connected open cone $\Omega \subset \mathbb{R}^2$ such that $\triangle u = |x_1|$ then there exists $a \in \mathbb{C}$ such that*

$$U(\theta) = \tfrac{1}{6}|\cos\theta|\cos(\theta)e^{i\theta} + \overline{a}e^{3i\theta} \tag{5-2}$$

for all $e^{i\theta} \in \Omega$ (in the rest of this section we identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$).

Proof. Let us write $v(x) = u(x) - \psi(x_1)$ with ψ as defined in (3-4); then v is a third-order homogeneous harmonic function in the connected open cone $\Omega \subset \mathbb{R}^2$. Thus there exists $a \in \mathbb{C}$ such that

$$v(x) = \Re(\overline{a}(x_1 + ix_2)^3) \quad \text{for all } x \in \Omega.$$

So we have

$$u(e^{i\theta}) = \tfrac{1}{6}|\cos\theta|^3 + \Re(\overline{a}e^{3i\theta}) \tag{5-3}$$

for all $e^{i\theta} \in \Omega$.

Differentiating (5-3) with respect to θ we obtain the desired equation. □

By the homogeneity of u it follows that

$$\{x \in \overline{\Omega} \mid u(x) = |\nabla u(x)| = 0\} = \{re^{i\theta} \in \overline{\Omega} \mid U(\theta) = 0,\ r > 0\}.$$

If $\Omega = \mathbb{R}^2\backslash\{0\}$ then, for u to be a solution to (5-1), U should be a periodic function with period 2π such that $U(\theta) \neq 0$ for all $\theta \in \mathbb{R}$ and if, in addition, u is a solution to (2-3) then we should have $\Re U(\theta) > 0$ for all $\theta \in \mathbb{R}$.

In the case that Ω is an open connected cone not equal to $\mathbb{R}^2\backslash\{0\}$, there exist $\theta_1, \theta_2 \in \mathbb{R}$ such that $\theta_1 < \theta_2 \leq \theta_1 + 2\pi$ and $\Omega = \{re^{i\theta} \mid r > 0,\ \theta_1 < \theta < \theta_2\}$. In this case, if u is a solution to (5-1) with $\Omega = \Omega(u)$, then U should satisfy $U(\theta_1) = U(\theta_2) = 0$ and $U(\theta) \neq 0$ for $\theta_1 < \theta < \theta_2$. If, in addition, u is a solution to (2-3) then we should have $\Re U(\theta) > 0$ for $\theta_1 < \theta < \theta_2$.

Let us define

$$V(\theta) = |\cos\theta|\cos(\theta)e^{2i\theta}. \tag{5-4}$$

It follows that

$$6e^{3i\theta}\overline{U}(\theta) = V(\theta) + 6a. \tag{5-5}$$

Lemma 18. *The function u is a solution of* (5-1) *with $\Omega = \mathbb{R}^2\backslash\{0\}$ if and only if $-6a \notin V(\mathbb{R})$.*

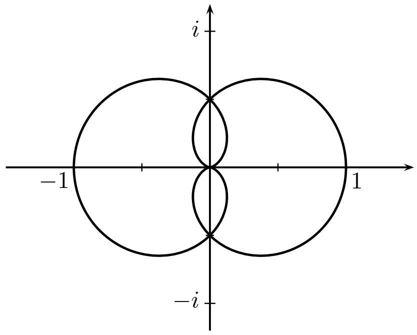


Figure 3. The set $V(\mathbb{R})$.

Proof. The function u is a solution of (5-1) with $\Omega = \mathbb{R}^2\backslash\{0\}$ if and only if U is 2π-periodic and $U(\theta) \neq 0$ for all $\theta \in \mathbb{R}$.

From (5-2) it follows that U is 2π-periodic and, by (5-5), it is clear that $U(\theta) \neq 0$ for all $\theta \in \mathbb{R}$ if and only if $-6a \notin V(\mathbb{R})$. ☐

From the definition of V in (5-4) it is clear that $B_1^c \subset (V(\mathbb{R}))^c$, so by Lemma 18 it follows that there are many solutions of (5-1) with $\Omega = \mathbb{R}^2\backslash\{0\}$.

Let us note that for a connected cone specified by θ_1 and θ_2, the solution with such a cone is unique. This follows from the fact that, because $U(\theta_1) = 0$, by (5-2) a is uniquely obtained and for this value of a the solution u is uniquely given by (5-3). Based on this observation, in the following we do not distinguish between a connected cone and the corresponding solution.

Lemma 19. *The function u is a solution of* (5-1) *with a connected open cone $\Omega \neq \mathbb{R}^2\backslash\{0\}$ if and only if one of the following cases hold*:

(i) $\theta_1 \notin \mathbb{Z}\pi + \left\{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ *and* $\theta_2 = \theta_1 + 2\pi$.

(ii) $\theta_1 \in \mathbb{Z}\pi + \frac{\pi}{2}$ *and* $\theta_2 = \theta_1 + \pi$.

(iii) $\theta_1 \in \mathbb{Z}\pi + \frac{\pi}{4}$ *and* $\theta_2 = \theta_1 + \frac{\pi}{2}$.

(iv) $\theta_1 \in \mathbb{Z}\pi + \frac{3\pi}{4}$ *and* $\theta_2 = \theta_1 + \frac{3\pi}{2}$.

Proof. Let us remember that we should have $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2 \leq \theta_1 + 2\pi$, $U(\theta_1) = U(\theta_2) = 0$ and $U(\theta) \neq 0$ for $\theta_1 < \theta < \theta_2$. It is possible to find all such θ_1 and θ_2 by algebraic computations, but for ease of presentation we resort to geometric arguments.

By (5-5), $U(\theta) = 0$ if and only if $-6a = V(\theta)$, hence we should have $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2 \leq \theta_1 + 2\pi$, $V(\theta_1) = V(\theta_2)$ and $V(\theta) \neq V(\theta_1)$ for $\theta_1 < \theta < \theta_2$. Thus we should find the smallest closed loops in the range graph of V. The range graph of V, i.e., the set $V(\mathbb{R})$ is depicted in Figure 3.

Then we have the following four cases:

(i) $-6a = V(\theta_1) \in V(\mathbb{R})\backslash\left\{0, \pm\frac{i}{2}\right\}$ with $\theta_1 \in \mathbb{R}\backslash\left(\mathbb{Z}\pi + \left\{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}\right)$ and the smallest loop is when $\theta_2 = \theta_1 + 2\pi$.

(ii) $-6a = V(\theta_1) = 0$ with $\theta_1 \in \mathbb{Z}\pi + \frac{\pi}{2}$ and the smallest loop is when $\theta_2 = \theta_1 + \pi$.

(iii) $-6a = V(\theta_1) \in \{\pm\frac{i}{2}\}$ with $\theta_1 \in \mathbb{Z}\pi + \frac{\pi}{4}$ and the smallest loop is when $\theta_2 = \theta_1 + \frac{\pi}{2}$.

(iv) $-6a = V(\theta_1) \in \{\pm\frac{i}{2}\}$ with $\theta_1 \in \mathbb{Z}\pi + \frac{3\pi}{4}$ and the smallest loop is when $\theta_2 = \theta_1 + \frac{3\pi}{2}$. □

There is some redundancy in the solutions specified in the previous lemma. In the following lemma we prove that if for two solutions the corresponding connected cones are rotations of each other by a multiple of π then the corresponding solutions are also rotated by the same angle.

Lemma 20. *Let $a, a' \in \mathbb{C}$ and let U, U' be the corresponding functions. If $n \in \mathbb{Z}$ and $\theta_0 \in \mathbb{R}$ are such that $U'(\theta_0 + n\pi) = U(\theta_0)$ then $U'(\theta + n\pi) = U(\theta)$ for all $\theta \in \mathbb{R}$.*

Proof. For any $n \in \mathbb{Z}$ and $\theta \in \mathbb{R}$ we have

$$\begin{aligned} U'(\theta + n\pi) &= \overline{a}' e^{3i(\theta+n\pi)} + \tfrac{1}{6}|\cos(\theta + n\pi)| \cos(\theta + n\pi) e^{i(\theta+n\pi)} \\ &= (-1)^n \overline{a}' e^{3i\theta} + \tfrac{1}{6}|\cos\theta| \cos(\theta) e^{i\theta} = ((-1)^n \overline{a'} - \overline{a}) e^{3i\theta} + U(\theta), \end{aligned}$$

from which the lemma follows because if $U'(\theta_0 + n\pi) = U(\theta_0)$ for some θ_0 then $(-1)^n \overline{a'} - \overline{a} = 0$, from which in turn it follows that $U'(\theta + n\pi) = U(\theta)$ for all θ. □

Corollary 21. *Let u and u' be solutions of* (5-1) *with $\Omega(u) = \{re^{i\theta} \mid \theta_1 < \theta < \theta_2,\ r > 0\}$ and $\Omega(u') = \{re^{i\theta} \mid \theta_1' < \theta < \theta_2',\ r > 0\}$, where $\theta_1 < \theta_2 \le \theta_1 + 2\pi$ and $\theta_1' < \theta_2' \le \theta_1' + 2\pi$. If there exists $n \in \mathbb{Z}$ such that $\theta_1' = \theta_1 + n\pi$ and $\theta_2' = \theta_2 + n\pi$, then $u'(e^{i(\theta+n\pi)}) = u(e^{i\theta})$ for $\theta_1 < \theta < \theta_2$.*

Proof. Let $U(\theta)$ correspond to $u(x)$ and $U'(\theta)$ to $u'(x)$. Then $U(\theta_1) = 0$ and $U'(\theta_1') = 0$. Thus $U(\theta_1) = U'(\theta_1') = U'(\theta_1 + n\pi)$. Now by Lemma 20 the corollary is proved. □

By this corollary we are able to remove some of the redundancies in Lemma 19, as stated in the following corollary:

Corollary 22. *The function u is a solution of* (5-1) *with a connected open cone $\Omega \ne \mathbb{R}^2 \backslash \{0\}$ if and only if one of the following cases hold*:

(i) $\theta_1 \in [0, 2\pi) \backslash \{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}\}$ *and $\theta_2 = \theta_1 + 2\pi$: the solutions corresponding to θ_1 in $[\pi, 2\pi) \backslash \{\frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}\}$ are equal to the solutions corresponding to $\theta_1 \in [0, \pi) \backslash \{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$ rotated by π, respectively.*

(ii) $\theta_1 \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$ *and $\theta_2 = \theta_1 + \pi$: the solution corresponding to $\theta_1 = \frac{3\pi}{2}$ is equal to the solution corresponding to $\theta_1 = \frac{\pi}{2}$ rotated by π,*

(iii) $\theta_1 \in \{\frac{\pi}{4}, \frac{5\pi}{4}\}$ *and $\theta_2 = \theta_1 + \frac{\pi}{2}$: the solution corresponding to $\theta_1 = \frac{5\pi}{4}$ is equal to the solution corresponding to $\theta_1 = \frac{\pi}{4}$ rotated by π.*

(iv) $\theta_1 \in \{\frac{3\pi}{4}, \frac{7\pi}{4}\}$ *and $\theta_2 = \theta_1 + \frac{3\pi}{2}$: the solution corresponding to $\theta_1 = \frac{7\pi}{4}$ is equal to the solution corresponding to $\theta_1 = \frac{3\pi}{4}$ rotated by π.*

By Lemma 18 we have obtained the solutions of (5-1) with $\Omega = \mathbb{R}^2 \backslash \{0\}$ and by Corollary 22 we have obtained all the solutions of (5-1) with a connected open cone $\Omega \ne \mathbb{R}^2 \backslash \{0\}$. Now we turn to finding the nonnegative solutions among these solutions.

To check the nonnegativity of a solution u, in the following lemma we write $u(e^{i\theta})$ in a closed form.

Lemma 23. *Let $\theta_1 < \theta_2 \le \theta_1 + 2\pi$ and let u be a solution to* (5-1) *in the cone corresponding to θ_1 and θ_2. Then we have*

$$6u(e^{i\theta}) = |\cos\theta|^3 - |\cos\theta_1|\cos(\theta_1)\cos(3\theta - 2\theta_1). \tag{5-6}$$

Proof. Because $U(\theta_1) = 0$, by (5-5) we have $6\overline{a} = -\overline{V}(\theta_1)$.

Now, by (5-3) we compute

$$\begin{aligned}6u(e^{i\theta}) = |\cos\theta|^3 + \Re(6\overline{a}e^{3i\theta}) &= |\cos\theta|^3 - \Re(\overline{V}(\theta_1)e^{3i\theta})\\ &= |\cos\theta|^3 - \Re(|\cos\theta_1|\cos(\theta_1)e^{-2i\theta_1}e^{3i\theta})\\ &= |\cos\theta|^3 - \Re(|\cos\theta_1|\cos(\theta_1)e^{(3\theta-2\theta_1)i})\\ &= |\cos\theta|^3 - |\cos\theta_1|\cos(\theta_1)\Re(e^{(3\theta-2\theta_1)i})\\ &= |\cos\theta|^3 - |\cos\theta_1|\cos(\theta_1)\cos(3\theta-2\theta_1),\end{aligned}$$

which proves (5-6). □

Lemma 24. *There exists no solution to the problem* (2-3) *with $\Omega = \{u > 0\} = \mathbb{R}^2\backslash\{0\}$.*

Proof. On the line segments $\{x_1 = 0\}\backslash\{0\}$, i.e., for $\theta = \pm\frac{\pi}{2}$, we have

$$\begin{aligned}6u(e^{\pm i\pi/2}) &= \left|\cos\left(\pm\tfrac{\pi}{2}\right)\right|^3 - |\cos\theta_1|\cos(\theta_1)\cos\left(\pm\tfrac{3\pi}{2} - 2\theta_1\right)\\ &= -|\cos\theta_1|\cos(\theta_1)\cos\left(\pm\tfrac{3\pi}{2} - 2\theta_1\right)\\ &= \pm|\cos\theta_1|\cos(\theta_1)\sin(2\theta_1).\end{aligned} \tag{5-7}$$

If $|\cos\theta_1|\cos(\theta_1)\sin(2\theta_1) = 0$ then $u(e^{\pm i\frac{\pi}{2}}) = 0$, which is in contradiction with $\Omega = \{u > 0\} = \mathbb{R}^2\backslash\{0\}$. If $|\cos\theta_1|\cos(\theta_1)\sin(2\theta_1) \neq 0$ then we can choose $\theta = \frac{\pi}{2}$ or $\theta = -\frac{\pi}{2}$ and obtain $u(e^{i\theta}) < 0$, which is again in contradiction with $\Omega = \{u > 0\} = \mathbb{R}^2\backslash\{0\}$. □

Lemma 25. *The function u is a solution of* (2-3) *with a connected open cone $\Omega \neq \mathbb{R}^2\backslash\{0\}$ if and only if one of the following cases hold*:

(i) *$\theta_1 \in \left\{\frac{\pi}{2}, \frac{3\pi}{2}\right\}$ and $\theta_2 = \theta_1 + \pi$: the solution corresponding to $\theta_1 = \frac{3\pi}{2}$ is equal to the solution corresponding to $\theta_1 = \frac{\pi}{2}$ rotated by π.*

(ii) *$\theta_1 \in \left\{\frac{\pi}{4}, \frac{5\pi}{4}\right\}$ and $\theta_2 = \theta_1 + \frac{\pi}{2}$: the solution corresponding to $\theta_1 = \frac{5\pi}{4}$ is equal to the solution corresponding to $\theta_1 = \frac{\pi}{4}$ rotated by π.*

Proof. We first show that the solutions given in parts (i) and (iv) of Corollary 22 are not nonnegative and then we show that the solutions given in parts (ii) and (iii) are nonnegative.

To prove the failure of nonnegativity of solutions given in part (i) of Corollary 22 we need only to consider $\theta_1 \in [0, \pi)\backslash\left\{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ with $\theta_2 = \theta_1 + 2\pi$ and, to prove the failure of nonnegativity of solutions given in part (iv), we need only to consider $\theta_1 = \frac{3\pi}{4}$ with $\theta_2 = \theta_1 + \frac{3\pi}{2}$.

For all these cases let us consider $\theta = \frac{3\pi}{2}$, then $\theta_1 < \theta < \theta_2$ and, by a similar computation as in (5-7), we obtain that

$$6u(e^{3\pi i/2}) = -|\cos\theta_1|\cos(\theta_1)\sin(2\theta_1).$$

Because for $\theta_1 \in [0,\pi)$ we have

$$|\cos\theta_1|\cos(\theta_1)\sin(2\theta_1) = 2|\cos\theta_1|\cos^2(\theta_1)\sin\theta_1 \geq 0,$$

this proves that the respective solutions take a nonpositive value at $\theta = \frac{3\pi}{2}$. If $u(e^{3\pi i/2}) < 0$ then u is not nonnegative. If $u(e^{3\pi i/2}) = 0$ and u was nonnegative then we would have $\partial_\theta u(e^{3\pi i/2}) = 0$, which is in contradiction with the connectedness of Ω.

To prove that the solutions given in Corollary 22(ii) are solutions of (2-3), we need only to consider the case when $\theta_1 = \frac{\pi}{2}$ and $\theta_2 = \theta_1 + \pi$. We compute

$$6u(e^{i\theta}) = |\cos\theta|^3 - \left|\cos\tfrac{\pi}{2}\right|\cos\left(\tfrac{\pi}{2}\right)\cos\left(3\theta - 2\left(\tfrac{\pi}{2}\right)\right) = |\cos\theta|^3 \tag{5-8}$$

and, because $|\cos\theta| > 0$ for $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$, we obtain that u is a solution of (2-3).

To prove that the solutions given in Corollary 22(iii) are solutions of (2-3), we need only to consider the case when $\theta_1 = \frac{\pi}{4}$ and $\theta_2 = \theta_1 + \frac{\pi}{2}$. We compute

$$6u(e^{i\theta}) = |\cos\theta|^3 - \left|\cos\tfrac{\pi}{4}\right|\cos\left(\tfrac{\pi}{4}\right)\cos\left(3\theta - \tfrac{\pi}{2}\right) = |\cos\theta|^3 - \tfrac{1}{2}\cos\left(3\theta - \tfrac{\pi}{2}\right) = |\cos\theta|^3 - \tfrac{1}{2}\sin(3\theta). \tag{5-9}$$

Let $\theta = \frac{\pi}{2} + \gamma$ for $-\frac{\pi}{4} < \gamma < \frac{\pi}{4}$; then

$$6u(e^{i(\pi/2+\gamma)}) = \left|\cos\left(\tfrac{\pi}{2} + \gamma\right)\right|^3 - \tfrac{1}{2}\sin\left(3\left(\tfrac{\pi}{2} + \gamma\right)\right) = |\sin\gamma|^3 + \tfrac{1}{2}\cos(3\gamma).$$

It follows that $6u(e^{i(\pi/2+\gamma)}) = 6u(e^{i(\pi/2-\gamma)})$, so we need only to consider $0 \leq \gamma < \frac{\pi}{4}$. For $0 \leq \gamma < \frac{\pi}{4}$ we have $\sin\gamma \geq 0$, thus

$$6u(e^{i(\pi/2+\gamma)}) = \sin^3\gamma + \tfrac{1}{2}\cos(3\gamma) = \tfrac{1}{2}\cos^3(\gamma)(\tan\gamma - 1)^2(2\tan\gamma + 1) > 0;$$

therefore we obtain that u is a solution of (2-3). ☐

Lemma 26. *In the original variable $x \in \mathbb{R}^2$, the only solutions of* (2-3) *with a connected open cone $\Omega \neq \mathbb{R}^2\backslash\{0\}$ are the following four solutions together with their noncoincidence cone Ω and their free boundary Γ:*

$$\begin{aligned}
&u(x) = u_{hs}(x), && \Omega = \{x_1 > 0\}, && \Gamma = \{x_1 = 0\};\\
&u(x) = u_{hs}(-x_1, x_2), && \Omega = \{x_1 < 0\}, && \Gamma = \{x_1 = 0\};\\
&u(x) = u_w(x), && \Omega = \{x_2 > |x_1|\}, && \Gamma = \{x_2 = |x_1|\};\\
&u(x) = u_w(x_1, -x_2), && \Omega = \{x_2 < -|x_1|\}, && \Gamma = \{x_2 = -|x_1|\}.
\end{aligned}$$

Proof. We compute the solutions given in Lemma 25 in the original variable.

For solutions given in Lemma 25(i), we only consider the case when $\theta_1 = \frac{\pi}{2}$ and $\theta_2 = \theta_1 + \pi$. We have

$$\left\{x = re^{i\theta} \;\middle|\; r > 0,\ \tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2}\right\} = \{x_1 < 0\}.$$

Now, for $x = re^{i\theta} \in \{x_1 < 0\}$, using the computation in (5-8) we compute

$$6u(x) = 6u(re^{i\theta}) = 6r^3u(e^{i\theta}) = r^3|\cos\theta|^3 = r^3|x_1/r|^3 = |x_1|^3 = (x_1^-)^3.$$

For solutions given in Lemma 25(ii) we only consider the case when $\theta_1 = \frac{\pi}{4}$ and $\theta_2 = \theta_1 + \frac{\pi}{2}$. We have

$$\{x = re^{i\theta} \mid r > 0,\ \tfrac{\pi}{4} < \theta < \tfrac{3\pi}{4}\} = \{x_2 > |x_1|\}.$$

Now, for $x = re^{i\theta} \in \{x_2 > |x_1|\}$, using the computation in (5-9) we compute

$$\begin{aligned} 6u(x) = 6u(re^{i\theta}) = 6r^3u(e^{i\theta}) &= r^3\big(|\cos\theta|^3 - \tfrac{1}{2}\sin(3\theta)\big) \\ &= r^3\big(|\cos\theta|^3 - \tfrac{1}{2}(3\cos^2(\theta)\sin\theta - \sin^3\theta)\big) \\ &= r^3\big(|x_1/r|^3 - \tfrac{1}{2}\big(3(x_1/r)^2x_2/r - (x_2/r)^3\big)\big) \\ &= |x_1|^3 - \tfrac{1}{2}(3x_1^2x_2 - x_2^3), \end{aligned}$$

which completes the proof of the lemma. □

Proof of Theorem 4. By Lemma 24 there exists no solution to the problem (2-3) with $\Omega = \{u > 0\} = \mathbb{R}^2\backslash\{0\}$.

So we are left only with solutions whose noncoincidence open cone Ω is a countable union of disjoint connected open cones. But, considering the only possible connected open cones as noncoincidence sets enumerated in Lemma 26, we come to the conclusion that, except for the solutions with connected cones, there exist two additional solutions, $u_w + u_w(x_1, -x_2)$ and $u_{hs} + u_{hs}(-x_1, x_2)$, each a combination of two solutions with connected open cones. □

Lemma 27. *We have*

$$W(1, u_{hs}) = \tfrac{\pi}{96} \quad \textit{and} \quad W(1, u_w) = \tfrac{1}{192}\big(\pi - \tfrac{8}{3}\big).$$

Proof. For any solution of (2-3) with connected open cone, we have, using (4-2),

$$\begin{aligned} W(1,u) = \int_{B_1} |x_1|u\,dx = \int_0^1\int_{\partial B_r} |x_1|u\,d\sigma(x)\,dr &= \int_0^1\int_{\partial B_1} |ry_1|u(ry)r\,d\sigma(y)\,dr \\ &= \int_0^1 r^5\,dr \int_{\partial B_1} |y_1|u(y)\,d\sigma(y) \\ &= \frac{1}{6}\int_{\theta_1}^{\theta_2} |\cos\theta|u(e^{i\theta})\,d\theta. \end{aligned}$$

For the half-space solution u_{hs}, we compute, using (5-8),

$$W(1, u_{hs}) = \frac{1}{36}\int_{\pi/2}^{3\pi/2} |\cos\theta|^4\,d\theta = \frac{1}{18}\int_0^{\pi/2} \cos^4\theta\,d\theta = \frac{\pi}{96}.$$

For the wedge solution u_w, we compute, using (5-9),

$$\begin{aligned} W(1,u_w) &= \frac{1}{36}\int_{\pi/4}^{3\pi/4} \big(|\cos\theta|^4 - \tfrac{1}{2}|\cos\theta|\sin(3\theta)\big)\,d\theta \\ &= \frac{1}{18}\int_{\pi/4}^{\pi/2} \cos^4\theta\,d\theta - \frac{1}{36}\int_{\pi/4}^{\pi/2} \cos(\theta)\sin(3\theta)\,d\theta = \tfrac{1}{192}\big(\pi - \tfrac{8}{3}\big), \end{aligned}$$

which completes the proof of the lemma. □

Corollary 28. *We have*

$$0 < W(1,u_w) = W(1,u_w(x_1,-x_2)) < W(1,u_w + u_w(x_1,-x_2)) = 2W(1,u_w) \\ < W(1,u_{hs}) = W(1,u_{hs}(-x_1,x_2)) < W(1,u_{hs}+u_{hs}(-x_1,x_2)) = 2W(1,u_{hs}).$$

Proof. The only inequality that is not clear is the inequality $2W(1,u_w) < W(1,u_{hs})$. But this is verified by the explicit values computed in the previous lemma. □

Corollary 29. *The set $\Gamma \cap \{x_1 = 0\}$ might be decomposed into four disjoint sets according to four possible values of the Weiss balanced energy. The closure of the set of points with a given energy w is a subset of the set of points with energy larger than or equal to w.*

Proof. Let $y \in \Gamma \cap \{x_1 = 0\}$; then, by the translation $u(x+y)$, we might assume that $y = 0$. Let $0 < \delta$ be such that $B_\delta \subset D$. Let us consider the family u_r for $0 < r < \frac{1}{6}\delta$. By Theorem 2 this family is uniformly bounded in $C^{1,1}(B_1)$. Thus there exists $r_j \to 0$ and $v \in C^{1,1}(B_1)$ such that $u_{r_j} \to v$ in $C^1(B_1)$. By Lemma 16, v is a nontrivial homogeneous global solution and $W(+0,u) = W(1,v)$. The possible values of $W(1,v)$ are only of the four values given in the previous corollary and this shows that the free boundary points $\Gamma \cap \{x_1 = 0\}$ divide into four disjoint sets depending on the Weiss balanced energy of the blowups at that point.

The last claim follows from the upper semicontinuity of $W(+0,x,u)$ stated in Lemma 14. □

For example, from Corollary 29 it follows that the set $\Gamma \cap \{x_1 = 0\} \cap \{W(+0,x,u) = 2W(1,u_{hs})\}$ is closed. Actually, at the end of Section 7 we will show that all points of $\Gamma \cap \{x_1 = 0\} \cap \{W(+0,x,u)$ in $\{W(1,u_w), 2W(1,u_w)\}\}$ are isolated points of $\Gamma \cap \{x_1 = 0\}$.

In the following lemma we obtain a lower bound for the homogeneous global solutions, which will be used in Lemma 32.

Lemma 30. *There exists a $c > 0$ such that for all homogeneous global solutions u we have*

$$u(x) \geq c d^2(x,\{u=0\})\big(d(x,\{u=0\}) + |x_1|\big) \quad \textit{for } x \in \mathbb{R}^2. \tag{5-10}$$

Proof. It is easy to see that we need to prove (5-10) for the cases when $u = u_w$ or $u = u_{hs}$.

In the case $u = u_{hs}$, for $x_1 \leq 0$ both sides of the inequality (5-10) are 0. For $x_1 > 0$ we have $d(x,\{u_{hs}=0\}) = x_1$, hence

$$u_{hs}(x) = \tfrac{1}{6}x_1^3 = \tfrac{1}{6}d^2(x,\{u_{hs}=0\})\big(\tfrac{1}{2}d(x,\{u_{hs}=0\}) + \tfrac{1}{2}x_1\big) \\ = \tfrac{1}{12}d^2(x,\{u_{hs}=0\})\big(d(x,\{u_{hs}=0\}) + x_1\big)$$

and this proves (5-10) for $u = u_{hs}$.

In the case $u = u_w$, for $x_2 < |x_1|$ both sides of the inequality are 0. Also, by the symmetry $u_w(x_1,x_2) = u_w(-x_1,x_2)$ we need only to consider the case $x_2 > x_1 > 0$.

For $x_2 > x_1 > 0$ it is easy to see that $d(x,\{u_w=0\}) = \frac{1}{\sqrt{2}}(x_2-x_1)$, thus for $x_2 > x_1 > 0$ we compute

$$\begin{aligned} u_w(x) = \tfrac{1}{6}x_1^3 + \tfrac{1}{12}x_2^3 - \tfrac{1}{4}x_1^2x_2 &= \tfrac{1}{12}(x_2-x_1)^2(2x_1+x_2) \\ &= \tfrac{1}{12}\big(\sqrt{2}d(x,\{u_w=0\})\big)^2\big(3x_1+\sqrt{2}d(x,\{u_w=0\})\big) \\ &\geq \tfrac{1}{6}\sqrt{2}d^2(x,\{u_w=0\})\big(d(x,\{u_w=0\})+x_1\big), \end{aligned}$$

which proves the desired inequality. □

In the next lemma we prove directional monotonicity type inequalities, which will be used in Lemma 37.

Lemma 31. *There exists a $C > 0$ such that $a\partial_\nu u_w - u_w \geq 0$ in $B_1 \cap \{(1+\epsilon)x_1 > x_2 > x_1 > 0\}$ if $\nu = e^{i(3\pi/4+\gamma)}$, $\epsilon > 0$, $-\frac{\pi}{2} < \gamma < \frac{\pi}{2}$ and $C(1/a+1)\epsilon \leq \cos\gamma$.*

Proof. For $x_2 > x_1 > 0$ we have

$$\begin{aligned} u_w(x) &= \tfrac{1}{6}x_1^3 + \tfrac{1}{12}x_2^3 - \tfrac{1}{4}x_1^2x_2 = \tfrac{1}{12}(x_2-x_1)^2(2x_1+x_2), \\ \partial_{x_1}u_w(x) &= \tfrac{1}{2}x_1^2 - \tfrac{1}{2}x_1x_2 = -\tfrac{1}{2}(x_2-x_1)x_1, \\ \partial_{x_2}u_w(x) &= \tfrac{1}{4}x_2^2 - \tfrac{1}{4}x_1^2 = \tfrac{1}{4}(x_2-x_1)(x_1+x_2). \end{aligned}$$

Thus we may compute, for $x_2 > x_1 > 0$,

$$\begin{aligned} a\partial_\nu u_w(x) - u_w(x) &= a\big(\nu_1\big(-\tfrac{1}{2}(x_2-x_1)x_1\big) + \nu_2\big(\tfrac{1}{4}(x_2-x_1)(x_1+x_2)\big)\big) - \tfrac{1}{12}(x_2-x_1)^2(2x_1+x_2) \\ &= \tfrac{1}{2}(x_2-x_1)\big(a\big(-\nu_1x_1 + \nu_2\big(\tfrac{1}{2}(x_1+x_2)\big)\big) - \tfrac{1}{6}(x_2-x_1)(2x_1+x_2)\big). \end{aligned} \tag{5-11}$$

Thus, to have $a\partial_\nu u_w(x) - u_w(x) \geq 0$ for $x \in \mathbb{R}^2$ satisfying $x_2 > x_1 > 0$ we should have

$$a\big(-\nu_1x_1 + \nu_2\big(\tfrac{1}{2}(x_1+x_2)\big)\big) \geq \tfrac{1}{6}(x_2-x_1)(2x_1+x_2)$$

and, rearranging this further, we get the equivalent inequality

$$\nu_2 - \nu_1 \geq \frac{1}{2x_1}(x_2-x_1)\Big(\frac{1}{3a}(2x_1+x_2) - \nu_2\Big).$$

Now, for $x \in B_1$ we have the bounds $x_1 < 1$ and $x_2 < 1$. Also, if $0 < x_1 < x_2$ then $x_2 - x_1 > 0$. So it is sufficient to have the inequality

$$\nu_2 - \nu_1 \geq \frac{1}{2x_1}(x_2-x_1)\Big(\frac{1}{a} - \nu_2\Big). \tag{5-12}$$

By $0 < x_1 < x_2 < (1+\epsilon)x_1$ we have $0 < (x_2-x_1)/x_1 < \epsilon$. Thus, if $1/a - \nu_2 > 0$ then we should have

$$\nu_2 - \nu_1 \geq \frac{\epsilon}{2}\Big(\frac{1}{a} - \nu_2\Big)$$

and if $1/a - \nu_2 \leq 0$ then we should have $\nu_2 - \nu_1 \geq 0$. Because $\nu_2 \geq -1$, for both cases it is sufficient to have

$$\nu_2 - \nu_1 \geq \frac{\epsilon}{2}\Big(\frac{1}{a} + 1\Big). \tag{5-13}$$

We compute

$$\nu_2 - \nu_1 = \sin\big(\tfrac{3\pi}{4}+\gamma\big) - \cos\big(\tfrac{3\pi}{4}+\gamma\big) = \sqrt{2}\cos\gamma. \tag{5-14}$$

From (5-13) and (5-14) it follows that it is sufficient to have

$$\cos\gamma \geq \frac{\sqrt{2}}{4}\Big(\frac{1}{a}+1\Big)\epsilon$$

and, taking $C \geq \frac{\sqrt{2}}{4}$, the second part is also proved. □

6. Uniqueness of blowup limits

Proof of Theorem 7. By Lemma 16 a blowup limit at the origin is a third-order homogeneous global solution.

By Theorem 4 we have six nontrivial homogeneous global solutions. Let us enumerate them by u^i for $i = 1, \dots, 6$.

Assume by contradiction that there exist $r_j \to 0$ and $\tilde{r}_j \to 0$ such that $u_{r_j} \to u^1$ and $u_{\tilde{r}_j} \to u^2$ in $C^1(B_1)$.

There exists $\epsilon > 0$ such that $\|u^i - u^1\|_{C(B_1)} > \epsilon$ for $i = 2, \dots, 6$.

Let us write $f(r) = \|u_r - u^1\|_{C(B_1)}$.

Because u is uniformly continuous in a neighbourhood of 0 we have that $f(r)$ is continuous for small enough $r > 0$. We have also $f(r_j) \to 0$ and $f(\tilde{r}_j) \to \|u^2 - u^1\|_{C(B_1)} > \epsilon$. Thus there exists $\hat{r}_j \to 0$ such that $f(\hat{r}_j) = \frac{1}{2}\epsilon$.

By Theorem 2, $u_{\hat{r}_j}$ is uniformly bounded in $C^{1,1}(B_1)$ for large j. Thus there exists a subsequence j_k such that $u_{\hat{r}_{j_k}}$ converges in C^1. By Lemma 16 the limit of $u_{\hat{r}_{j_k}}$ is a third-order nontrivial homogeneous global solution. This is in contradiction with $f(\hat{r}_{j_k}) = \frac{1}{2}\epsilon$ and the choice of ϵ. □

7. Convergence of the free boundary to the free boundary of the blowup limit

In the following lemma, roughly speaking, we prove two inclusions. First, if u is close to a nontrivial homogeneous global solution u_0 then, for x far from $\{u_0 = 0\}$, we have $u(x) > 0$. Second, if u is close to a solution u_0 then, for x far from $\{u_0 > 0\}$, we have $x \in \{u = 0\}^\circ$.

Lemma 32. *There exists $c > 0$ such that if u_0 is a nontrivial homogeneous global solution and u is a solution in B_1, then we have*

$$\big\{x \in B_1 \;\big|\; cd^2(x, \{u_0 = 0\})\big(d(x, \{u_0 = 0\}) + |x_1|\big) > \|u - u_0\|_{L^\infty(B_1)}\big\} \subset \{u > 0\}; \tag{7-1}$$

here $\{u_0 = 0\} = \{x \in \mathbb{R}^2 \mid u_0(x) = 0\}$ and $\{u > 0\} = \{x \in B_1 \mid u(x) > 0\}$.

If u_0 and u are solutions in B_1 and

$$\|u - u_0\|_{L^\infty(B_1)} < c,$$

then

$$\big\{x \in B_{1/2} \;\big|\; cd^2(x, \{u_0 > 0\})\big(d(x, \{u_0 > 0\}) + |x_1|\big) > \|u - u_0\|_{L^\infty(B_1)}\big\} \subset \{u = 0\}^\circ; \tag{7-2}$$

here $\{u_0 = 0\} = \{x \in B_1 \mid u_0(x) = 0\}$ and $\{u = 0\} = \{x \in B_1 \mid u(x) = 0\}$.

Proof. Assume u_0 is a nontrivial homogeneous global solution and u is a solution in B_1. Using Lemma 30 for $x \in B_1$ we compute

$$\begin{aligned} u(x) = u_0(x) + u(x) - u_0(x) &\geq u_0(x) - \|u - u_0\|_{L^\infty(B_1)} \\ &\geq c_1 d^2(x, \{u_0 = 0\})\big(d(x, \{u_0 = 0\}) + |x_1|\big) - \|u - u_0\|_{L^\infty(B_1)}; \end{aligned}$$

here c_1 is the constant in Lemma 30. So, if

$$\|u - u_0\|_{L^\infty(B_1)} < \tfrac{1}{2} c_1 d^2(x, \{u_0 = 0\})\big(d(x, \{u_0 = 0\}) + |x_1|\big)$$

then

$$u(x) > \tfrac{1}{2} c_1 d^2(x, \{u_0 = 0\})\big(d(x, \{u_0 = 0\}) + |x_1|\big)$$

and this proves (7-1) with $0 < c \leq \frac{1}{2} c_1$.

Assume u_0 and u are solutions in B_1. By Theorem 15 there exists $c_2 > 0$ such that, if $y \in B_1$, $u(y) > 0$ and $B_r(y) \Subset B_1$, then we have

$$\sup_{\{u>0\}\cap\partial B_r(y)} u \geq u(y) + c_2 r^2 (r + |y_1|).$$

Thus, if $y \in B_1$, $u(y) > 0$, $B_r(y) \Subset \{u_0 = 0\} \cap B_1$ and $c_2 r^2 (r + |y_1|) > \|u - u_0\|_{L^\infty(B_1)}$, then we have

$$\begin{aligned} 0 = \sup_{\{u>0\}\cap\partial B_r(y)} u_0 &= \sup_{\{u>0\}\cap\partial B_r(y)} (u - (u - u_0)) \\ &\geq \sup_{\{u>0\}\cap\partial B_r(y)} u - \|u - u_0\|_{L^\infty(B_1)} \\ &\geq u(y) + c_2 r^2 (r + |y_1|) - \|u - u_0\|_{L^\infty(B_1)} \\ &\geq c_2 r^2 (r + |y_1|) - \|u - u_0\|_{L^\infty(B_1)}; \end{aligned}$$

a contradiction. Thus, if $y \in B_1$, $B_r(y) \Subset \{u_0 = 0\} \cap B_1$ and $c_2 r^2 (r + |y_1|) > \|u - u_0\|_{L^\infty(B_1)}$, then $u(y) = 0$.

For $y \in (\{u_0 = 0\} \cap B_1)^\circ$, setting $r = \frac{1}{2} d\big(y, (\{u_0 = 0\} \cap B_1)^c\big)$ it follows that if

$$\tfrac{1}{4} c_2 d^2\big(y, (\{u_0 = 0\} \cap B_1)^c\big)\big(\tfrac{1}{2} d\big(y, (\{u_0 = 0\} \cap B_1)^c\big) + |y_1|\big) > \|u - u_0\|_{L^\infty(B_1)}$$

then $u(y) = 0$. This proves that

$$\big\{x \in B_1 \,\big|\, \tfrac{1}{8} c_2 d^2\big(x, (\{u_0 = 0\} \cap B_1)^c\big)\big(d\big(x, (\{u_0 = 0\} \cap B_1)^c\big) + |x_1|\big) > \|u - u_0\|_{L^\infty(B_1)}\big\} \subset \{u = 0\}.$$

By the continuity of $d\big(x, (\{u_0 = 0\} \cap B_1)^c\big)$ as a function of x it follows that

$$\big\{x \in B_1 \,\big|\, \tfrac{1}{8} c_2 d^2\big(x, (\{u_0 = 0\} \cap B_1)^c\big)\big(d\big(x, (\{u_0 = 0\} \cap B_1)^c\big) + |x_1|\big) > \|u - u_0\|_{L^\infty(B_1)}\big\} \subset \{u = 0\}^\circ. \quad (7\text{-}3)$$

Let $x \in B_{1/2}$; then we compute

$$d\big(x, (\{u_0 = 0\} \cap B_1)^c\big) = d(x, \{u_0 > 0\} \cup B_1^c) = \min\big(d(x, \{u_0 > 0\}), d(x, B_1^c)\big) \geq \min\big(d(x, \{u_0 > 0\}), \tfrac{1}{2}\big),$$

so we have

$$
\begin{aligned}
d^2\big(x,(\{u_0=0\}\cap B_1)^c\big)\big(d\big(x,(\{u_0=0\}\cap B_1)^c\big)+|x_1|\big)
&= \min\Big(d^2(x,\{u_0>0\})\big(d(x,\{u_0>0\})+|x_1|\big), \big(\tfrac{1}{2}\big)^2\big(\tfrac{1}{2}+|x_1|\big)\Big)\\
&\geq \min\big(d^2(x,\{u_0>0\})\big(d(x,\{u_0>0\})+|x_1|\big), \tfrac{1}{8}\big). \qquad (7\text{-}4)
\end{aligned}
$$

So, by (7-3) and (7-4), if

$$\|u-u_0\|_{L^\infty(B_1)} < \tfrac{1}{64}c_2$$

then

$$\big\{x\in B_{1/2} \;\big|\; \tfrac{1}{8}c_2 d^2(x,\{u_0>0\})\big(d(x,\{u_0>0\})+|x_1|\big) > \|u-u_0\|_{L^\infty(B_1)}\big\} \subset \{u=0\}^\circ \qquad (7\text{-}5)$$

and, by choosing $0<c\leq \frac{1}{64}c_2$, this finishes the proof of the lemma. □

By the inclusions proved in the previous lemma, in the following lemma we show that for u a solution and u_0 a nontrivial homogeneous global solution, if u is close enough to u_0 then the free boundary of u is in a quantitatively specified neighbourhood of the free boundary of u_0.

Lemma 33. *There exists $c>0$ such that, if u is a solution in B_1 and u_0 is a nontrivial homogeneous global solution, then if*

$$\|u-u_0\|_{L^\infty(B_1)} < c \qquad (7\text{-}6)$$

we have

$$\Gamma\cap B_{1/2} \subset \big\{cd^2(x,\Gamma_{u_0})(d(x,\Gamma_{u_0})+|x_1|) \leq \|u-u_0\|_{L^\infty(B_1)}\big\}.$$

Proof. If $u=u_0$ in B_1 then the claim is obvious, so we assume that $u_0\neq u$ in B_1.

Assume there exists $x\in\Gamma\cap B_{1/2}$ such that

$$cd^2(x,\Gamma_{u_0})(d(x,\Gamma_{u_0})+|x_1|) > \|u-u_0\|_{L^\infty(B_1)};$$

here $c>0$ is as in Lemma 32.

Then, because

$$d(x,\Gamma_{u_0}) = \max\big(d(x,\{u_0=0\}), d(x,\{u_0>0\})\big),$$

we should have either

$$cd^2(x,\{u_0=0\})\big(d(x,\{u_0=0\})+|x_1|\big) > \|u-u_0\|_{L^\infty(B_1)} \qquad (7\text{-}7)$$

or

$$cd^2(x,\{u_0>0\})\big(d(x,\{u_0>0\})+|x_1|\big) > \|u-u_0\|_{L^\infty(B_1)}. \qquad (7\text{-}8)$$

In the case when (7-7) holds then, by (7-1), we obtain that $u(x)>0$, which is in contradiction with $x\in\Gamma$.

In the case when (7-8) holds then, because also (7-6) holds by (7-2), we obtain that $x\in\{u=0\}^\circ$, which is in contradiction with $x\in\Gamma$ and this finishes the proof of the lemma. □

Lemma 34. *There exists $c>0$ such that if u_0 is a nontrivial homogeneous global solution, u is a solution in D, $0\in D$ and $0\in\Gamma$ then, for $x\in\Gamma$ such that $B_{4|x|}\subset D$ and*

$$\|u_{4|x|}-u_0\|_{L^\infty(B_1)}<c,$$

we have

$$cd^2(x,\Gamma_{u_0})(d(x,\Gamma_{u_0})+|x_1|)\le|x|^3\|u_{4|x|}-u_0\|_{L^\infty(B_1)}.$$

Proof. Let c be as in Lemma 32.

Let $r>0$ and assume

$$\|u_r-u_0\|_{L^\infty(B_1)}<c;$$

then, by Lemma 33, we have

$$\Gamma_{u_r}\cap B_{1/2}\subset\big\{cd^2(y,\Gamma_{u_0})(d(y,\Gamma_{u_0})+|y_1|)\le\|u_r-u_0\|_{L^\infty(B_1)}\big\}.$$

Then, because Γ_{u_0} is a cone and $\Gamma_u\cap B_{r/2}=r(\Gamma_{u_r}\cap B_{1/2})$, we obtain

$$\begin{aligned}\Gamma_u\cap B_{r/2}&\subset\big\{ry\in B_{r/2}\;\big|\;cd^2(y,\Gamma_{u_0})\big(d(y,\Gamma_{u_0})+|y_1|\big)\le\|u_r-u_0\|_{L^\infty(B_1)}\big\}\\&=\Big\{x\in B_{r/2}\;\Big|\;cd^2\Big(\frac{x}{r},\Gamma_{u_0}\Big)\Big(d\Big(\frac{x}{r},\Gamma_{u_0}\Big)+\Big|\frac{x_1}{r}\Big|\Big)\le\|u_r-u_0\|_{L^\infty(B_1)}\Big\}\\&=\big\{x\in B_{r/2}\;\big|\;cd^2(x,\Gamma_{u_0})(d(x,\Gamma_{u_0})+|x_1|)\le r^3\|u_r-u_0\|_{L^\infty(B_1)}\big\}.\end{aligned}$$

For those $x\in\Gamma_u$ such that $B_{4|x|}\subset D$, we may consider $r=4|x|$.

So, if

$$\|u_{4|x|}-u_0\|_{L^\infty(B_1)}<c$$

then, because $x\in\Gamma_u\cap B_{2|x|}$, we have

$$cd^2(x,\Gamma_{u_0})\big(d(x,\Gamma_{u_0})+|x_1|\big)\le 4^3|x|^3\|u_{4|x|}-u_0\|_{L^\infty(B_1)}.$$ □

Proof of Theorem 8. Let us consider the case $W(+0,u)=W(1,u_w)$ with the blowup limit u_w. Then for $x\in\{x_1>0,\ x_2>-x_1\}$ we have $d(x,\Gamma_{u_w})=\frac{\sqrt{2}}{2}|x_2-x_1|$ and, for $x\in\{x_1>0,\ x_2\le -x_1\}$, we have $d(x,\Gamma_{u_w})=|x|\ge\frac{\sqrt{2}}{2}|x_2-x_1|$. Thus we compute, for $x_1>0$,

$$d(x,\Gamma_{u_w})+|x_1|\ge\tfrac{1}{2}\sqrt{2}|x_2-x_1|+|x_1|\ge c_1|x|. \tag{7-9}$$

By symmetry we obtain the same inequality for $x_1<0$.

Now, by Lemma 34 we obtain the inequality (2-7). For the remaining cases, when $W(+0,u)$ is in $\{W(1,u_w),2W(1,u_w)\}$, we can compute similarly.

In the cases when $W(+0,u)\in\{W(1,u_{hs}),2W(1,u_{hs})\}$ we have $\Gamma_{u_0}=\{x_1=0\}$ and $d(x,\Gamma_{u_0})=|x_1|$, so (2-8) follows immediately from Lemma 34. □

Corollary 35. *Let u be a solution in D; then the points of*

$$\Gamma\cap\{x_1=0\}\cap\big\{W(+0,x,u)\in\{W(1,u_w),2W(1,u_w)\}\big\}$$

are isolated points of $\Gamma\cap\{x_1=0\}$ (in the topology of $\{x_1=0\}$).

Proof. Assume $W(+0,u) \in \{W(1,u_w), 2W(1,u_w)\}$; then, by (2-7), the free boundary should converge to the free boundary of the blowup limit tangentially. But this is not the case if the origin is not an isolated point of $\Gamma \cap \{x_1 = 0\}$. □

8. Convergence of the normal of the free boundary to the normal of the free boundary of the blowup limit at regular points

In the following lemma we prove a nondegeneracy type result for $u - a\partial_\nu u$ far from the degeneracy line $\{x_1 = 0\}$.

Lemma 36. *If u is a solution in D, $y \in \Omega$, $B_r(y) \Subset D \cap \{x_1 \geq \frac{1}{16}\}$ and $u(y) - \frac{1}{32}\partial_\nu u(y) > 0$, then we have*

$$\tfrac{1}{128} r^2 \leq \sup_{\Omega \cap \partial B_r(y)} (u - a\partial_\nu u).$$

Proof. Let y and r be as in the statement of the theorem.

We define, for $a > 0$ and $c > 0$,

$$h(x) = u(x) - a\partial_\nu u(x) - (u(y) - a\partial_\nu u(y)) - c|x-y|^2.$$

We compute

$$\triangle h(x) = |x_1| - a\nu_1 x_1/|x_1| - 4c \geq \tfrac{1}{16} - a - 4c \quad \text{in } \Omega \cap \{x_1 \geq \tfrac{1}{16}\},$$

so if we choose $a = \frac{1}{32}$ and $c = \frac{1}{128}$ then we have

$$\triangle h \geq 0 \quad \text{in } \Omega \cap \{x_1 \geq \tfrac{1}{16}\}. \tag{8-1}$$

Also we have

$$h(y) = 0. \tag{8-2}$$

For $x \in \Gamma$ we have $u(x) - \frac{1}{32}\partial_\nu u(x) = 0$, thus if $u(y) - \frac{1}{32}\partial_\nu u(y) > 0$ then we have

$$h(x) = -\big(u(y) - \tfrac{1}{32}\partial_\nu u(y)\big) - \tfrac{1}{128}|x-y|^2 < 0 \quad \text{on } \Gamma. \tag{8-3}$$

Because $B_r(y) \subset \{x_1 \geq \frac{1}{16}\}$, by (8-1) we have that h is subharmonic in the domain $\Omega \cap B_r(y)$. Applying the maximum principle for the domain $\Omega \cap B_r(y)$ and the subharmonic function h, we have

$$h(y) \leq \sup_{\partial(\Omega \cap B_r(y))} h. \tag{8-4}$$

By (8-2) and (8-4), we obtain

$$0 \leq \sup_{\partial(\Omega \cap B_r(y))} h. \tag{8-5}$$

Because

$$\partial(\Omega \cap B_r(y)) = (\partial\Omega \cap B_r(y)) \cup (\Omega \cap \partial B_r(y)),$$

by (8-3) and (8-5) we obtain

$$0 \leq \sup_{\Omega \cap \partial B_r(y)} h. \tag{8-6}$$

By the definition of h, from (8-6) we get the inequality

$$u(y)-\tfrac{1}{32}\partial_\nu u(y)+\tfrac{1}{128}r^2\le \sup_{\Omega\cap\partial B_r(y)}\bigl(u-\tfrac{1}{32}\partial_\nu u\bigr) \tag{8-7}$$

and this proves the lemma. □

Let ν_w be the normal to $\Gamma_{u_w}\cap\{x_1>0\}$ pointing into $\{u_w>0\}$, i.e.,

$$\nu_w=\tfrac{1}{\sqrt{2}}(-1,1).$$

In the following lemma we prove a crucial directional monotonicity result, which will be used in the proof of the convergence of normals.

Lemma 37. *There exists $c>0$ such that, if u is a solution in B_1, $x_u\in\Gamma_u\cap\partial B_{1/4}\cap\{x_1>0\}$, $\nu\in\partial B_1$ and $r>0$ are such that*

$$\|u-u_w\|_{C^1(B_1)}^{1/2}+r\le c\nu\cdot\nu_w,$$

then

$$\tfrac{1}{32}\partial_\nu u-u\ge 0\quad in\ \ \Omega\cap B_r(x_u).$$

Proof. We have

$$\{x_{u_w}\}=\Gamma_{u_w}\cap\partial B_{1/4}\cap\{x_1>0\},\quad \text{where}\quad x_{u_w}=\tfrac{\sqrt{2}}{8}(1,1).$$

Step 1. In this step we show that there exists $C_1>0$ such that

$$|x_u-x_{u_w}|\le C_1\|u-u_w\|_{L^\infty(B_1)}^{1/2}. \tag{8-8}$$

By Lemma 33 there exists $c>0$ such that if $\|u-u_w\|_{L^\infty(B_1)}<c$ then

$$\Gamma_u\cap B_{1/2}\subset\bigl\{c(d(x,\Gamma_{u_w}))^2(d(x,\Gamma_{u_w})+|x_1|)\le\|u-u_w\|_{L^\infty(B_1)}\bigr\}. \tag{8-9}$$

We have $x_u\in\Gamma_u\cap\partial B_{1/4}\cap\{x_1>0\}$; thus, by (8-9),

$$c(d(x_u,\Gamma_{u_w}))^2(d(x_u,\Gamma_{u_w})+|x_{u,1}|)\le\|u-u_w\|_{L^\infty(B_1)}. \tag{8-10}$$

As in (7-9) there exists $c_1>0$ such that

$$d(x_u,\Gamma_{u_w})+|x_{u,1}|\ge c_1|x_u|=\tfrac14 c_1. \tag{8-11}$$

Also, because $x_u\in\partial B_{1/4}\cap\{x_1>0\}$ there exists $C_2>0$ such that

$$|x_u-x_{u_w}|\le C_2 d(x_u,\Gamma_{u_w}). \tag{8-12}$$

Now, by (8-10), (8-11) and (8-12), it follows that there exists $C_3>0$ such that

$$|x_u-x_{u_w}|\le C_3\|u-u_w\|_{L^\infty(B_1)}^{1/2}. \tag{8-13}$$

Step 2. In this step we show that there exists $\delta>0$ such that if

$$\|u-u_w\|_{L^\infty(B_1)}<\delta\quad\text{and}\ \ 0<r<\tfrac{1}{48}$$

then, for $x \in \Omega \cap B_{1/48}(x_u)$, if $u(x) - \frac{1}{32}\partial_\nu u(x) > 0$ we have

$$\tfrac{1}{128} r^2 \le \sup_{\Omega \cap \partial B_r(x)} \bigl(u - \tfrac{1}{32}\partial_\nu u\bigr).$$

By Step 1, if

$$C_3 \|u - u_w\|^{1/2}_{L^\infty(B_1)} < \tfrac{1}{48}$$

then $|x_u - x_{u_w}| < \frac{1}{48}$. Thus $x_{u,1} > x_{u_w,1} - \frac{1}{48}$ and

$$B_{1/48}(x_u) \subset \bigl\{x_1 > x_{u_w,1} - \tfrac{1}{48} - \tfrac{1}{48}\bigr\} = \{x_1 > x_{u_w,1} - \tfrac{1}{24}\}$$

and, for $x \in B_{1/48}(x_u)$, we have

$$B_{1/48}(x) \subset \bigl\{x_1 > x_{u_w,1} - \tfrac{1}{24} - \tfrac{1}{48}\bigr\} = \bigl\{x_1 > x_{u_w,1} - \tfrac{1}{16}\bigr\} = \bigl\{x_1 > \tfrac{\sqrt{2}}{8} - \tfrac{1}{16}\bigr\} \subset \{x_1 > \tfrac{1}{8} - \tfrac{1}{16}\} = \bigl\{x_1 > \tfrac{1}{16}\bigr\}.$$

Now, by Lemma 36, if

$$0 < r < \tfrac{1}{48},$$

$x \in \Omega \cap B_{1/48}(x_u)$ and $u(x) - \frac{1}{32}\partial_\nu u(x) > 0$, then we have

$$\tfrac{1}{128} r^2 \le \sup_{\Omega \cap \partial B_r(x)} \bigl(u - \tfrac{1}{32}\partial_\nu u\bigr).$$

Step 3. In this step we show that there exists $C_4 > 0$ such that $\frac{1}{32}\partial_\nu u_w - u_w \ge 0$ in $B_\eta(x_{u_w})$ if $0 < \eta < \frac{1}{16}$, $\nu \in \partial B_1$ and $C_4 \eta \le \nu \cdot \nu_w$.

Assume $x \in B_\eta(x_{u_w})$ with $0 < \eta < \frac{1}{16}$. Then

$$x_1 > x_{u_w,1} - \eta > x_{u_w,1} - \tfrac{1}{16} = \tfrac{\sqrt{2}}{8} - \tfrac{1}{16} > \tfrac{1}{8} - \tfrac{1}{16} = \tfrac{1}{16}$$

and

$$\begin{aligned} \frac{x_2}{x_1} = 1 + \frac{x_2 - x_1}{x_1} \le 1 + \frac{|x_2 - x_1|}{x_1} &< 1 + 16|x_2 - x_1| = 1 + 16\sqrt{2}\, d(x, \{x_2 = x_1\}) \\ &\le 1 + 16\sqrt{2}|x - x_{u_w}| \le 1 + 16\sqrt{2}\eta; \end{aligned}$$

hence by Lemma 31 we have $\frac{1}{32}\partial_\nu u_w(x) - u_w(x) \ge 0$ if $\nu \in \partial B_1$ and

$$C\bigl(\tfrac{1}{1/32} + 1\bigr)(16\sqrt{2}\eta) \le \nu \cdot \nu_w$$

with $C > 0$ as in Lemma 31.

Step 4. In this step we show that there exists $\delta_1 > 0$ and $C_5 > 0$ such that, if

$$\|u - u_w\|_{L^\infty(B_1)} < \delta_1, \quad 0 < r < \tfrac{1}{48}, \quad 0 < r_1 < \tfrac{1}{48}, \tag{8-14}$$

$$\nu \in \partial B_1, \quad C_4\bigl(r + r_1 + C_3\|u - u_w\|^{1/2}_{L^\infty(B_1)}\bigr) \le \nu \cdot \nu_w, \tag{8-15}$$

$$C_5 \|u - u_w\|^{1/2}_{C^1(B_1)} < r, \tag{8-16}$$

then

$$u - \tfrac{1}{32}\partial_\nu u \le 0 \quad \text{in } \Omega \cap B_{r_1}(x_u). \tag{8-17}$$

By Step 1 there exists $0 < \delta_1 < \delta$ such that if

$$\|u - u_w\|_{L^\infty(B_1)} < \delta_1 \tag{8-18}$$

then

$$|x_u - x_{u_w}| < \tfrac{1}{48}. \tag{8-19}$$

Let

$$0 < r < \tfrac{1}{48} \quad \text{and} \quad 0 < r_1 < \tfrac{1}{48}. \tag{8-20}$$

Assume now that both (8-18) and (8-20) hold.

We define

$$\eta = r + r_1 + |x_u - x_{u_w}|;$$

then by (8-19) and (8-20) we have

$$0 < \eta < \tfrac{1}{16}. \tag{8-21}$$

By Step 2 for $x \in \Omega \cap B_{r_1}(x_u)$, if $u(x) - \frac{1}{32}\partial_\nu u(x) > 0$ then

$$\tfrac{1}{128} r^2 \le \sup_{\Omega \cap \partial B_r(x)} \big(u - \tfrac{1}{32}\partial_\nu u\big). \tag{8-22}$$

By (8-21) and Step 3 we have $\frac{1}{32}\partial_\nu u_w - u_w \ge 0$ in $B_\eta(x_{u_w})$ if

$$\nu \in \partial B_1 \quad \text{and} \quad C_4 \eta \le \nu \cdot \nu_w. \tag{8-23}$$

Assume now that (8-23) holds.

We have

$$B_r(x) \subset B_{r+|x-x_u|}(x_u) \subset B_{r+|x-x_u|+|x_u-x_{u_w}|}(x_{u_w}) \subset B_{r+r_1+|x_u-x_{u_w}|}(x_{u_w}) \subset B_\eta(x_{u_w}).$$

We compute

$$\begin{aligned}\sup_{\Omega \cap \partial B_r(x)} \big(u - \tfrac{1}{32}\partial_\nu u\big) &\le \sup_{\Omega \cap \partial B_r(x)} \big(u_w - \tfrac{1}{32}\partial_\nu u_w\big) + \sup_{\Omega \cap \partial B_r(x)} \big(u - \tfrac{1}{32}\partial_\nu u - \big(u_w - \tfrac{1}{32}\partial_\nu u_w\big)\big) \\ &\le C_6 \|u - u_w\|_{C^1(B_1)}.\end{aligned}$$

Therefore, by (8-22), if

$$\tfrac{1}{128} r^2 > C_6 \|u - u_w\|_{C^1(B_1)}$$

then

$$u - \tfrac{1}{32}\partial_\nu u \le 0 \quad \text{in } \Omega \cap B_{r_1}(x_u).$$

Step 5. In this step we finish the proof of the lemma.

Choosing

$$r = 2C_5 \|u - u_w\|_{C^1(B_1)}^{1/2},$$

(8-16) holds. Noticing that $\nu \cdot \nu_w \le 1$ we obtain that, by choosing $c > 0$ small enough, if

$$\nu \in \partial B_1, \quad \|u - u_w\|_{C^1(B_1)}^{1/2} + r_1 \le c\nu \cdot \nu_w$$

holds then (8-14) and (8-15) hold and thus, by Step 4, (8-17) holds and this proves the lemma. □

For $0 \le \delta < 1$ let us define the open cone

$$C_\delta = \{x \in \mathbb{R}^2 \mid x \cdot \nu_w > \delta |x|\}.$$

Corollary 38. *If u is a solution in B_1, $x \in \Gamma \cap \partial B_{1/4} \cap \{x_1 > 0\}$, $0 < \delta < 1$ and $r > 0$ are such that*

$$\|u - u_w\|_{C^1(B_1)}^{1/2} + r \le c\delta$$

with $c > 0$ as in Lemma 37, then

$$B_r(x) \cap (x + C_\delta) \subset \{u > 0\} \quad \text{and} \quad B_r(x) \cap (x - C_\delta) \subset \{u = 0\}. \tag{8-24}$$

Proof. By Lemma 37 and the definition of C_δ we have that, for all $\nu \in C_\delta$,

$$\partial_\nu u \ge 0 \quad \text{in } B_r(x_u). \tag{8-25}$$

From (8-30), because $u \ge 0$,

$$z \in B_r(x) \quad \text{and} \quad u(z) = 0 \quad \Longrightarrow \quad B_r(x) \cap (z - C_\delta) \subset \{u = 0\}. \tag{8-26}$$

In particular, because $u(x) = 0$ we have

$$B_r(x) \cap (x - C_\delta) \subset \{u = 0\}.$$

Now assume there exists $y \in B_r(x) \cap (x + C_\delta)$ such that $u(y) = 0$. By (8-26) we have that $u = 0$ in $B_r(x) \cap (y - C_\delta)$. From $y \in x + C_\delta$ it follows that $x \in y - C_\delta$, thus x is in the interior of $B_r(x) \cap (y - C_\delta)$, where we have shown that $u = 0$ and this contradicts $x \in \Gamma$. □

It is easy to see that, for the cone C'_δ conjugate to the cone C_δ, we have

$$C'_\delta = \{x \in \mathbb{R}^2 \mid x \cdot y \ge 0 \text{ for all } y \in C_\delta\} = \overline{C}_{\sqrt{1-\delta^2}}. \tag{8-27}$$

Theorem 39. *There exists $C_1 > 0$ such that, if u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular point with blowup limit u_w, then there exists $\epsilon > 0$ such that all points of $\Gamma \cap \{x_1 > 0\} \cap B_\epsilon$ are usual (for $x_1 > 0$ the force term is nondegenerate) regular free boundary points and*

$$|n(x) - \nu_w| \le C_1 \|u_{4|x|} - u_w\|_{C^1(B_1)}^{1/2} \tag{8-28}$$

for $x \in \Gamma \cap \{x_1 > 0\} \cap B_\epsilon$, where $n(x)$ is the normal to Γ at x, pointing into Ω.

Proof. If there exists $r > 0$ such that $u = u_w$ in B_r then the claim of the theorem holds trivially. So we might assume that for all $r > 0$ we have $u \ne u_w$ in B_r.

Let $x \in \Gamma \cap \{x_1 > 0\} \cap B_1$. By the uniqueness of the blowup limit and Theorem 2 we have that $u_{4|x|} \to u_w$ in $C^1(B_1)$ as $x \to 0$. Thus there exists $\epsilon > 0$ such that for $|x| < \epsilon$ we have

$$\|u_{4|x|} - u_w\|_{C^1(B_1)} < \left(\frac{c}{2}\right)^2 \tag{8-29}$$

with $c > 0$ as in Lemma 37.

Let $y = \frac{1}{4}x/|x|$. Then $y \in \Gamma_{u_{4|x|}} \cap \partial B_{1/4} \cap \{x_1 > 0\}$. By (8-29), if we choose

$$\delta = \frac{2}{c}\|u_{4|x|} - u_w\|^{1/2}_{C^1(B_1)} \tag{8-30}$$

then $0 < \delta < 1$.

Also let us set

$$r = \|u_{4|x|} - u_w\|^{1/2}_{C^1(B_1)}. \tag{8-31}$$

Then, by (8-30) and (8-31) we have

$$\|u_{4|x|} - u_w\|^{1/2}_{C^1(B_1)} + r = c\delta \tag{8-32}$$

and consequently, by Corollary 38, we have

$$B_r(y) \cap (y + C_\delta) \subset \{u_{4|x|} > 0\} \quad \text{and} \quad B_r(y) \cap (y - C_\delta) \subset \{u_{4|x|} = 0\}. \tag{8-33}$$

From (8-33) it follows that

$$B_{4|x|r}(x) \cap (x + C_\delta) \subset \{u > 0\} \quad \text{and} \quad B_{4|x|r}(x) \cap (x - C_\delta) \subset \{u = 0\}. \tag{8-34}$$

Now, if x is a singular free boundary point then the blowup limit is a nonzero homogeneous quadratic polynomial. But, by (8-34), this polynomial should be equal to 0 in $-C_\delta$, which brings us to contradiction. Thus all points of $\Gamma \cap \{x_1 > 0\} \cap B_\epsilon$ are regular points.

Now assume $|x| < \epsilon$; then, because x is a regular point, Γ has a normal at this point. Let $n(x)$ be the normal to Γ pointing into Ω. From (8-34) it follows that $n(x) \in C'_\delta$. Now, by (8-27), we have

$$n(x) \in \overline{C}_{\sqrt{1-\delta^2}},$$

so

$$n(x) \cdot \nu_w \geq \sqrt{1-\delta^2}.$$

We compute

$$|n(x) - \nu_w|^2 = 2 - 2n(x) \cdot \nu_w \leq 2 - 2\sqrt{1-\delta^2} = \frac{2\delta^2}{1 + \sqrt{1-\delta^2}} \leq 2\delta^2 \tag{8-35}$$

and (8-28) follows from (8-30) and (8-35). □

9. Free boundary as a graph near regular points

The following two lemmas will be used in Lemma 42.

Lemma 40. *If u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular free boundary point with blowup limit u_w, then there exists an $\epsilon > 0$ such that $u(0,t) > 0$ for $0 < t < \epsilon$ and $(0,t) \in \{u = 0\}^\circ$ for $-\epsilon < t < 0$.*

Proof. Let $x = (0,t) \in B_\epsilon$, $0 < t < \epsilon$, then we compute

$$\begin{aligned} d^2\big(\tfrac{1}{2}x/|x|, \{u_w = 0\}\big)\big(d\big(\tfrac{1}{2}x/|x|, \{u_w = 0\}\big) + \big|\tfrac{1}{2}x_1/|x|\big|\big) &= d^3\big(\tfrac{1}{2}x/|x|, \{u_w = 0\}\big) \\ &= d^3\big(\tfrac{1}{2}e_2, \{u_w = 0\}\big) = \big(\tfrac{\sqrt{2}}{4}\big)^3. \end{aligned} \tag{9-1}$$

For small enough ϵ, if $|x| < \epsilon$ then

$$\|u_{2|x|} - u_w\|_{L^\infty(B_1)} < c\left(\tfrac{\sqrt{2}}{4}\right)^3 \tag{9-2}$$

with c as in Lemma 32. Thus, by (9-1), (9-2) and (7-1), we have $u_{2|x|}\left(\frac{1}{2}x/|x|\right) > 0$, so $u(x) > 0$.

Let $x = (0, t) \in B_\epsilon$, $-\epsilon < t < 0$, then we compute

$$\begin{aligned} d^2\left(\tfrac{1}{4}x/|x|, \{u_w > 0\}\right)\left(d\left(\tfrac{1}{4}x/|x|, \{u_w > 0\}\right) + \left|\tfrac{1}{4}x_1/|x|\right|\right) &= d^3\left(\tfrac{1}{4}x/|x|, \{u_w > 0\}\right) \\ &= d^3\left(-\tfrac{1}{4}e_2, \{u_w > 0\}\right) = \tfrac{1}{4^3}. \end{aligned} \tag{9-3}$$

For small enough ϵ, if $|x| < \epsilon$ then

$$\|u_{4|x|} - u_w\|_{L^\infty(B_1)} < \tfrac{1}{4^3}c. \tag{9-4}$$

Thus, by (9-3), (9-4) and (7-2), we have $x/(4|x|) \in \{u_{4|x|} = 0\}^\circ$, so $x \in \{u = 0\}^\circ$. □

Lemma 41. *If u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular free boundary point with blowup limit u_w, then there exists an $\epsilon > 0$ such that for every $0 < x_1 < \frac{1}{4}\epsilon$ there exists a unique x_2 such that $x = (x_1, x_2) \in \Gamma \cap B_\epsilon$ and, for $(x_1, t) \in B_\epsilon$, we have $u(x_1, t) > 0$ if $t > x_2$ and $(x_1, t) \in \{u = 0\}^\circ$ if $t < x_2$.*

Proof. First we show that there exists $\epsilon > 0$ such that for all $0 < x_1 < \frac{1}{4}\epsilon$ there exists x_2 such that $(x_1, x_2) \in \Gamma \cap B_\epsilon$.

Let $\epsilon > 0$, to be chosen later. Let $0 < x_1 < \frac{1}{4}\epsilon$; then we compute

$$\left|\left(x_1, \tfrac{3}{4}\epsilon\right)\right|^2 < \left(\tfrac{1}{4}\epsilon\right)^2 + \left(\tfrac{3}{4}\epsilon\right)^2 = \tfrac{10}{16}\epsilon^2 < \epsilon^2;$$

thus $\left(x_1, \frac{3}{4}\epsilon\right) \in B_\epsilon$. We compute

$$d\left(\left(x_1/\epsilon, \tfrac{3}{4}\right), \{u_w = 0\}\right) = \tfrac{\sqrt{2}}{2}\left(\tfrac{3}{4} - x_1/\epsilon\right) \geq \tfrac{\sqrt{2}}{2}\left(\tfrac{3}{4} - \tfrac{1}{4}\right) = \tfrac{\sqrt{2}}{4}$$

and

$$d^2\left(\left(x_1/\epsilon, \tfrac{3}{4}\right), \{u_w = 0\}\right)\left(d\left(\left(x_1/\epsilon, \tfrac{3}{4}\right), \{u_w = 0\}\right) + |x_1/\epsilon|\right) \geq d^3\left(\left(x_1/\epsilon, \tfrac{3}{4}\right), \{u_w = 0\}\right) \geq \left(\tfrac{\sqrt{2}}{4}\right)^3.$$

Thus, if ϵ is small enough that

$$\|u_\epsilon - u_w\|_{L^\infty(B_1)} < c\left(\tfrac{\sqrt{2}}{4}\right)^3$$

with c as in Lemma 32, then by (7-1) we obtain that

$$u_\epsilon\left(x_1/\epsilon, \tfrac{3}{4}\right) > 0$$

and therefore

$$u\left(x_1, \tfrac{3}{4}\epsilon\right) > 0. \tag{9-5}$$

Let $0 < x_1 < \frac{1}{4}\epsilon$; then we compute

$$\left|\left(x_1, -\tfrac{1}{4}\epsilon\right)\right|^2 < \left(\tfrac{1}{4}\epsilon\right)^2 + \left(\tfrac{1}{4}\epsilon\right)^2 = \left(\tfrac{\sqrt{2}}{4}\epsilon\right)^2 < \left(\tfrac{1}{2}\epsilon\right)^2,$$

thus $\big(x_1, -\frac{1}{4}\epsilon\big) \in B_{\epsilon/2} \subset B_\epsilon$.

We compute

$$d\big((x_1/\epsilon, -\tfrac{1}{4}), \{u_w > 0\}\big) \geq \tfrac{1}{4}$$

and

$$d^2\big((x_1/\epsilon, -\tfrac{1}{4}), \{u_w > 0\}\big)\big(d\big((x_1/\epsilon, -\tfrac{1}{4}), \{u_w > 0\}\big) + |x_1/\epsilon|\big) \geq \tfrac{1}{4^3}.$$

Thus, if ϵ is small enough that

$$\|u_\epsilon - u_w\|_{L^\infty(B_1)} < \tfrac{1}{4^3}c,$$

then by (7-2) we obtain that

$$\big(x_1/\epsilon, -\tfrac{1}{4}\big) \in \{u_\epsilon = 0\}^\circ$$

and therefore

$$\big(x_1, -\tfrac{1}{4}\epsilon\big) \in \{u = 0\}^\circ. \tag{9-6}$$

From (9-5), (9-6) and the continuity of u it follows that there exists $-\frac{1}{4}\epsilon < x_2 < \frac{3}{4}\epsilon$ such that $(x_1, x_2) \in \Gamma$. This finishes the proof of the existence of x_2.

By Corollary 38 there exists $c > 0$ such that, if $y \in \Gamma_u \cap \{y_1 > 0\}$ and

$$\|u_{4|y|} - u_w\|_{C^1(B_1)} \leq \big(\tfrac{1}{4}c\big)^2,$$

then

$$B_{c|y|}(y) \cap (y + C_{1/2}) \subset \{u > 0\} \quad \text{and} \quad B_{c|y|}(y) \cap (y - C_{1/2}) \subset \{u = 0\}. \tag{9-7}$$

Now let ϵ be small enough that $\sigma_1(4\epsilon) \leq \big(\frac{1}{4}c\big)^2$. Then (9-7) holds for $y \in \Gamma_u \cap B_\epsilon \cap \{y_1 > 0\}$.

Because $x = (x_1, x_2) \in \Gamma_u \cap B_\epsilon \cap \{x_1 > 0\}$, by (9-7) we have

$$B_{c|x|}(x) \cap (x + C_{1/2}) \subset \{u > 0\} \quad \text{and} \quad B_{c|x|}(x) \cap (x - C_{1/2}) \subset \{u = 0\}. \tag{9-8}$$

Assume there exists $(x_1, t) \in B_\epsilon$ such that $t > x_2$ and $u(t, x_2) = 0$. Let t^* be the infimum of such t, i.e.,

$$t^* = \inf\{t > x_2 \mid (x_1, t) \in B_\epsilon \ \text{and} \ u(t, x_2) = 0\}.$$

From the first inclusion in (9-8) we have that $t^* > x_2$. Thus for $x_2 < s < t^*$ we have $u(x_1, s) > 0$, therefore (x_1, t^*) is on the boundary of $\{u > 0\}$. We obtain that $(x_1, t^*) \in \Gamma_u$. But now, because $(x_1, t^*) \in \Gamma_u \cap B_\epsilon \cap \{x_1 > 0\}$, by the second inclusion in (9-7) at the point (x_1, t^*) we come to a contradiction.

Now assume that there exists $(x_1, t) \in B_\epsilon$ such that $t < x_2$ and $(t, x_2) \in \overline{\{u > 0\}}$. Let t^* be the supremum of such t, i.e.,

$$t^* = \sup\{t < x_2 \mid (x_1, t) \in B_\epsilon \cap \overline{\{u > 0\}}\}.$$

From the second inclusion in (9-8) we have that $t^* < x_2$. Thus for $t^* < s < x_2$ we have $(x_1, s) \in \{u = 0\}^\circ$, therefore $(x_1, t^*) \in \Gamma_u$. But now, because $(x_1, t^*) \in \Gamma_u \cap B_\epsilon \cap \{x_1 > 0\}$, by the first inclusion in (9-7) at the point (x_1, t^*) we come to a contradiction. □

In the following lemma we prove that near to regular points the free boundary is a continuous graph.

Lemma 42. *If u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular free boundary point with blowup limit u_w, then there exists an $\epsilon > 0$ and $\gamma \in C\big(\big[0, \frac{1}{4}\epsilon\big)\big)$ such that $\gamma(0) = 0$, we have $(x_1, \gamma(x_1)) \in B_\epsilon$ for $0 < x_1 < \frac{1}{4}\epsilon$, and*

$$\{u = 0\} \cap B_\epsilon \cap \big\{0 \le x_1 < \tfrac{1}{4}\epsilon\big\} = \big\{x \in B_\epsilon \;\big|\; 0 \le x_1 < \tfrac{1}{4}\epsilon,\ x_2 \le \gamma(x_1)\big\}. \tag{9-9}$$

Proof. By Lemma 41 there exists an $\epsilon > 0$ such that, for every $0 < x_1 < \frac{1}{4}\epsilon$, there exists a unique x_2 such that $x = (x_1, x_2) \in \Gamma \cap B_\epsilon$; let us define $\gamma(x_1) = x_2$. Let us also define $\gamma(0) = 0$.

Then, by Lemmas 40 and 41, we have (9-9).

Now let us show that γ is continuous. Assume there exists $0 \le y < \frac{1}{4}\epsilon$ such that γ is discontinuous at y. Then there exists $x_j \to y$ such that $\gamma(x_j) \to z$ and either $z > \gamma(y)$ or $z < \gamma(y)$.

In the case $z > \gamma(y)$ we have $u(y, z) > 0$, which is in contradiction with $u(x_j, \gamma(x_j)) = 0$ and the continuity of u.

In the case $z < \gamma(y)$ we have $(y, z) \in \{u = 0\}^\circ$, which is in contradiction with $(x_j, \gamma(x_j)) \in \Gamma$. □

In the following lemma we formulate the convergence of the free boundary in terms of the function γ.

Lemma 43. *There exists $C_1 > 0$ and $C_2 > 0$ such that, if u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular free boundary point with blowup limit u_w, then, with $\epsilon > 0$ and γ as in Lemma 42, we have*

$$|\gamma(x_1) - x_1| \le C_1\big(\sigma_0(C_2|x_1|)\big)^{1/2}|x_1| \quad \textit{for } 0 < x_1 < \tfrac{1}{4}\epsilon,$$

where σ_0 is as defined in (2-6).

Proof. By Theorem 8 we have

$$d(x, \Gamma_{u_w}) \le C_1\big(\sigma_0(C_2|x|)\big)^{1/2}|x|.$$

For $x_1 > 0$ we estimate

$$d(x, \Gamma_{u_w}) \ge \tfrac{\sqrt{2}}{2}|x_2 - x_1|;$$

thus

$$\begin{aligned}|\gamma(x_1) - x_1| \le C_3\big(\sigma_0(C_2|x|)\big)^{1/2}|x| &\le C_4\big(\sigma_0(C_2|x|)\big)^{1/2}\big(|\gamma(x_1)| + |x_1|\big)\\ &\le C_4\big(\sigma_0(C_2|x|)\big)^{1/2}\big(|\gamma(x_1) - x_1| + 2|x_1|\big).\end{aligned} \tag{9-10}$$

By the continuity of γ at 0 we have that $\gamma(x_1) \to \gamma(0) = 0$ as $x_1 \to 0$. Hence $|x| \le C_5\big(|\gamma(x_1)| + |x_1|\big) \to 0$ as $x_1 \to 0$. From this convergence we obtain $\sigma_0(C_2|x|) \to 0$ as $x_1 \to 0$.

Thus, from (9-10) it follows that

$$|\gamma(x_1) - x_1| \le C_6\big(\sigma_0(C_2|x|)\big)^{1/2}|x_1|. \tag{9-11}$$

In turn, from (9-11) it follows that

$$\begin{aligned}|x| \le C_5\big(|\gamma(x_1)| + |x_1|\big) \le C_5\big(|\gamma(x_1) - x_1| + 2|x_1|\big) &\le C_5\big(C_6\big(\sigma_0(C_2|x|)\big)^{1/2}|x_1| + 2|x_1|\big)\\ &= C_5\big(C_6\big(\sigma_0(C_2|x|)\big)^{1/2} + 2\big)|x_1|\\ &\le C_7|x_1|.\end{aligned} \tag{9-12}$$

Now, by (9-11) and (9-12) the lemma is proved. □

In the following lemma we formulate the convergence of the normals in terms of the function γ.

Lemma 44. *There exists $C_1 > 0$ and $C_2 > 0$ such that, if u is a solution in D, $0 \in D$ and $0 \in \Gamma$ is a regular free boundary point with blowup limit u_w, and $\epsilon > 0$ and γ are as in Lemma 42, then we have $\gamma \in C^1\big(0, \frac{1}{4}\epsilon\big)$ and*

$$|\gamma'(x_1) - 1| \le C_1\big(\sigma_1(C_2|x_1|)\big)^{1/2},$$

where σ_1 is as defined in (2-6).

Proof. By Theorem 39, for small enough $\epsilon > 0$ all points of $\Gamma \cap \{x_1 > 0\} \cap B_\epsilon$ are usual regular points. Let $0 < x_1 < \frac{1}{4}\epsilon$. Hence (see [Petrosyan et al. 2012]) Γ is a C^1 curve in a neighbourhood of $(x_1, \gamma(x_1))$. From (8-28) it follows that for small enough ϵ and $|x| < \epsilon$ we have $n(x) \notin \{-e_1, e_1\}$. It follows that $\gamma'(x_1)$ exists and

$$n(x) = \frac{(-\gamma'(x_1), 1)}{\sqrt{1 + (\gamma'(x_1))^2}}.$$

From here it follows that there exists $C > 0$ such that for $n(x)$ close enough to ν_w we have

$$|\gamma'(x_1) - 1| \le C|n(x) - \nu_w|. \tag{9-13}$$

Now, by (8-28) and (9-13) we obtain

$$|\gamma'(x_1) - 1| \le C_2\|u_{4|x|} - u_w\|^{1/2}_{C^1(B_1)}. \tag{9-14}$$

By (9-12) together with the definition of σ_1 and (9-14), the lemma is proved. □

Proof of Theorem 9. This follows from Lemmas 42, 43 and 44 and the symmetry of the problem with respect to the line $\{x_1 = 0\}$. □

In the case when 0 is a regular point but with $u_w(x_1, -x_2)$ as the blowup limit, we consider the even reflection $\tilde{u}(x_1, x_2) = u(x_1, -x_2)$, apply Theorem 9 to $\tilde{u}$ and obtain that the free boundary of u is a graph with properties as in Theorem 9 but reflected with respect to the line $\{x_2 = 0\}$.

By the following two lemmas we prove that if $W(+0, u) = 2W(1, u_w)$ then u might be decomposed into the sum of two functions each having 0 as a regular point.

Lemma 45. *If u is a solution in D, $0 \in D$, $0 \in \Gamma$ and $W(+0, u) = 2W(1, u_w)$, then there exists an $\epsilon > 0$ such that $u(x_1, 0) = 0$ for $|x_1| < \epsilon$.*

Proof. Let $u_0 = u_w + u_w(x_1, -x_2)$. We have

$$d\big(\pm\tfrac{1}{4}e_1, \{u_0 > 0\}\big) = \tfrac{\sqrt{2}}{8}.$$

We compute

$$d^2\big(\pm\tfrac{1}{4}e_1, \{u_0 > 0\}\big)\big(d\big(\pm\tfrac{1}{4}e_1, \{u_0 > 0\}\big) + \tfrac{1}{4}\big) = \big(\tfrac{\sqrt{2}}{8}\big)^2\big(\tfrac{\sqrt{2}}{8} + \tfrac{1}{4}\big).$$

Now, if $|x_1| > 0$ is small enough that

$$\|u_{4|x_1|} - u_0\|_{L^\infty(B_1)} < c\big(\tfrac{\sqrt{2}}{8}\big)^2\big(\tfrac{\sqrt{2}}{8} + \tfrac{1}{4}\big)$$

with c as in Lemma 32 then, by (7-2), we have $u_{4|x_1|}\big(\pm\frac{1}{4}e_1\big)=0$. Thus $u(x_1,0)=0$. □

Lemma 46. *If u is a solution in D, $0\in D$, $0\in\Gamma$ and $W(+0,u)=2W(1,u_w)$, then there exists an $\epsilon>0$ such that $u_+=\chi_{\{x_2>0\}}u$ and $u_-=\chi_{\{x_2<0\}}u$ are solutions in B_ϵ. We have $W(+0,u_\pm)=W(1,u_w)$, the blowup limit of u_+ is u_w and the blowup limit of u_- is $u_w(x_1,-x_2)$.*

Proof. By Lemma 45 there exists an $\epsilon>0$ such that $u(x_1,0)=0$ for $|x_1|<\epsilon$.

Because $u\geq 0$, $u\in C^1_{\mathrm{loc}}(D)$ and $u(x_1,0)=0$ for $|x_1|<\epsilon$, it follows that $\nabla u(x_1,0)=0$ for $|x_1|<\epsilon$.

From this it follows that u_+ and u_- are solutions in B_ϵ. We have $u_r(x)\to u_w+u_w(x_1,-x_2)$ in $C^1(B_1)$ as $r\to 0$. Thus $\chi_{\{x_2>0\}}u_r\to u_w$ in $C^1(B_1)$ and $u_{+,r}(x)=r^{-3}\chi_{\{x_2>0\}}(rx)u(rx)=\chi_{\{x_2>0\}}u_r(x)$; hence $u_{+,r}(x)\to u_w$ in $C^1(B_1)$ and

$$W(+0,u_+)=\lim_{r\to+0}W(r,u_+)=\lim_{r\to+0}W(1,u_{+,r})=W(1,u_w).$$

We argue similarly for u_-. □

In the case $W(+0,u)=2W(1,u_w)$, by Lemma 46 and Theorem 9 it follows that the free boundary near to 0 is the union of two graphs, one graph as in Theorem 9 and the other a graph with properties as in Theorem 9 but reflected with respect to the line $\{x_2=0\}$.

10. An irregularity result for the free boundary near degenerate points

Lemma 47. *Let u be a solution in D with $0\in D$. Suppose also that there exists $\delta>0$ such that $B_\delta\subset D$, $\partial_{x_2}u\leq 0$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$, $\Gamma\cap B_\delta\cap\{x_1=0,\ x_2>0\}\neq\varnothing$ and $B_\delta\cap\{x_1>0,\ x_2>0\}\subset\Omega$; then $u=u_{hs}$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$.*

Proof. For ease of notation let us write $v=-\partial_{x_2}u$. We have that v is harmonic in Ω and $v\geq 0$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$.

Assume $y\in\Gamma\cap B_\delta\cap\{x_1=0,\ x_2>0\}$, then by the optimal growth (Theorem 11) we have $\partial_{x_1}v(y)=0$. For small enough $r>0$ we have $B_r(re_1+y)\subset\Omega$. Now, because v is nonnegative and harmonic in $B_r(re_1+y)$ and $\partial_{x_1}v(y)=0$, by Hopf's lemma we conclude that $v=0$ in $B_r(re_1+x)$. Because v is harmonic in Ω we obtain that $v=0$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$. Hence $u=u(x_1)$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$. By this and the assumption $\Gamma\cap B_\delta\cap\{x_1=0,\ x_2>0\}\neq\varnothing$ the claim follows. □

Lemma 48. *Let u be a solution in D with $0\in D$. Suppose also that there exists $\delta>0$ such that $B_\delta\subset D$, $\partial_{x_2}u\leq 0$ in $B_\delta\cap\{x_1>0,\ x_2>0\}$, and there exists $\rho\in C\big(\big[0,\frac{1}{2}\delta\big)\big)\cap C^1\big(\big[0,\frac{1}{2}\delta\big)\big)$ such that $\rho(0)=\rho'(+0)=0$, $\rho>0$ in $\big(0,\frac{1}{2}\delta\big)$, ρ is convex and*

$$\Omega\cap B_\delta\cap\big\{x_1>0,\ 0<x_2<\tfrac{1}{2}\delta\big\}=B_\delta\cap\big\{0<x_2<\tfrac{1}{2}\delta,\ \rho(x_2)<x_1\big\};\tag{10-1}$$

then for every $q>1$ there exist $c>0$ and $t_0>0$ such that

$$\rho(t)\geq ct^q\quad\text{and}\quad\rho'(t)\geq ct^{q-1}\quad\text{for } 0<t<t_0.\tag{10-2}$$

Proof. Again, for ease of notation let us write $v=-\partial_{x_2}u$. The proof is divided into multiple steps.

Step 1. In this step we show that $v > 0$ in $B_\delta \cap \{0 < x_2 < \frac{1}{2}\delta,\ \rho(x_2) < x_1\}$.

If there is $x \in B_\delta \cap \{0 < x_2 < \frac{1}{2}\delta,\ \rho(x_2) < x_1\}$ such that $v(x) = 0$ then, because v is harmonic and nonnegative in $B_\delta \cap \{0 < x_2 < \frac{1}{2}\delta,\ \rho(x_2) < x_1\}$, it follows that $v = 0$ in $B_\delta \cap \{0 < x_2 < \frac{1}{2}\delta,\ \rho(x_2) < x_1\}$, but then because $u(\rho(t), t) = 0$ for $0 < t < \frac{1}{2}\delta$ we come to contradiction with (10-1).

Step 2. In this step we show that for each $q > 1$ and $\eta > \big(\tan(\pi/(2q))\big)^{-1}$ there exist $c_1 > 0$ (depending on u) and $t_1 > 0$ such that

$$v(x_t) \geq c_1 t^{2q} \quad \text{for } 0 < t < t_1, \tag{10-3}$$

where

$$x_t = (\eta t, t) \in \Omega.$$

Let $q > 1$ and

$$\alpha_q = \frac{\pi}{2q}.$$

Because $\rho'(+0) = 0$ there exists $t_q > 0$ such that $\rho(t) < t/\tan\alpha_q$ for $0 < t < t_q$.

Let us denote

$$r_q = \frac{t_q}{\tan\alpha_q}.$$

It follows that

$$\Omega_q = \{x = re^{i\theta} \mid 0 < r < r_q,\ 0 < \theta < \alpha_q\} \subset \Omega.$$

Let us define the function

$$v_q(x) = r^{2q}\sin(2q\theta) \quad \text{for } x = re^{i\theta} \in \Omega_q.$$

We have

$$\partial\big(\tfrac{1}{2}\Omega_q\big) = S_q \cup A_q,$$

where

$$S_q = \big\{x = re^{i\theta} \,\big|\, 0 \leq r < \tfrac{1}{2}r_q,\ \theta \in \{0, \alpha_q\}\big\}$$

and

$$A_q = \big\{x = re^{i\theta} \,\big|\, r = \tfrac{1}{2}r_q,\ 0 \leq \theta \leq \alpha_q\big\}.$$

Let $a = \frac{1}{2}r_q e_1$ and $b = \frac{1}{2}r_q e^{i\alpha_q}$ be the endpoints of the arc A_q. We have $b \in \Omega$, hence $v(b) > 0$. Either $v(a) > 0$ or $v(a) = 0$ and, by Hopf's lemma, we have $\partial_{x_2} v(a) > 0$. Also we have $v > 0$ on $A_q \backslash \{a, b\}$.

Thus there exists $\epsilon > 0$ such that

$$\epsilon v_q \leq v \quad \text{on } A_q. \tag{10-4}$$

We have $v_q = 0$ and $v \geq 0$ on S_q, thus

$$\epsilon v_q \leq v \quad \text{on } S_q. \tag{10-5}$$

Putting (10-4) and (10-5) together we have

$$\epsilon v_q \leq v \quad \text{on } \partial\big(\tfrac{1}{2}\Omega_q\big).$$

Now, by the maximum principle, we obtain that

$$\epsilon v_q \le v \quad \text{in } \tfrac{1}{2}\Omega_q. \tag{10-6}$$

We compute $|x_t| = \sqrt{1+\eta^2}t$, so for

$$0 < t < \frac{1}{2}\frac{r_q}{\sqrt{1+\eta^2}}$$

we have $|x_t| < \frac{1}{2}r_q$; also we compute

$$\frac{x_{t,2}}{x_{t,1}} = \frac{1}{\eta} < \tan\alpha_q,$$

thus we have

$$x_t \in \frac{1}{2}\Omega_q \quad \text{for } 0 < t < \frac{1}{2}\frac{r_q}{\sqrt{1+\eta^2}}. \tag{10-7}$$

Now, by (10-6) and (10-7) we have

$$v(x_t) \ge \epsilon v_q(x_t) = \epsilon |x_t|^{2q} \sin\Bigl(2q \arctan\frac{1}{\eta}\Bigr) = c_1 t^{2q} \quad \text{for } 0 < t < \frac{1}{2}\frac{r_q}{\sqrt{1+\eta^2}},$$

where

$$c_1 = \epsilon(1+\eta^2)^q \sin\Bigl(2q\arctan\frac{1}{\eta}\Bigr) > 0.$$

Step 3. In this step we show that there exists $c_2 > 0$ (independent of u) and $t_2 > 0$ such that if

$$0 < t < t_2 \quad \text{and} \quad \eta < 1$$

then there exists $y_t = (\rho(y_{t,2}), y_{t,2}) \in \Gamma$ with $0 < y_{t,2} < t_q$ such that

$$d_t = |y_t - x_t| = d(\Gamma, x_t)$$

and

$$\partial_{n(y_t)} v(y_t) \ge \frac{c_2}{d_t} v(x_t). \tag{10-8}$$

Here $n(y)$ is the normal to Γ at y, pointing into Ω.

Let

$$\Pi_q = \{0 < x_1 < r_q,\ 0 < x_2 < t_q\};$$

then we have

$$\Gamma_q = \Gamma \cap \Pi_q = \{(\rho(t), t) \mid 0 < t < t_q\}.$$

One may see that

$$d(x_t, \partial\Pi_q) = \min\bigl\{\eta t,\ r_q - \eta t,\ t,\ t_q - t\bigr\} = \eta t \tag{10-9}$$

if

$$t < \min\Bigl(\frac{r_q}{2\eta}, \frac{t_q}{1+\eta}\Bigr) \quad \text{and} \quad \eta < 1.$$

Because $\eta > (\tan\alpha_q)^{-1}$ and $0 < t < t_q$, we have that $\rho(t) < t/\tan\alpha_q < \eta t$. Also we have $\rho(t) > 0$, thus

$$d(x_t, (\rho(t), t)) = \eta t - \rho(t) < \eta t.$$

Now, because $(\rho(t), t) \in \Gamma_q$ we have

$$d(x_t, \Gamma) < \eta t. \tag{10-10}$$

By (10-9) and (10-10) there exists $y_t \in \Gamma_q$ such that

$$d_t = |y_t - x_t| = d(\Gamma, x_t). \tag{10-11}$$

Because

$$d(x_t, \partial\Pi_q) = \eta t > d(\Gamma, x_t) = d_t,$$

we have

$$B_{d_t}(x_t) \subset \Pi_q \subset \Omega.$$

Because $y_t \in \partial B_{d_t}(x_t)$, by the quantitative Hopf lemma (see [Han and Lin 2011]) there exists $c_2 > 0$ (independent of u and t) such that (10-8) holds.

Step 4. In this step we show that

$$\partial_{n(y)} v(y) = -n_2(y) y_1 \quad \text{for } y \in \Gamma_q. \tag{10-12}$$

By the equation $\triangle u = |x_1| \chi_{\{u>0\}}$ and the smoothness of the free boundary Γ_q, i.e., smoothness of ρ, it follows that in a neighbourhood of $y \in \Gamma_q$ we have

$$\triangle v = -n_2 |x_1| \mathcal{H}^1 \llcorner \Gamma. \tag{10-13}$$

From (10-1) and (10-13), the equation (10-12) follows.

Step 5. In this step we show that for $0 < t < t_2$ we have

$$y_{t,2} < (1+\eta)t. \tag{10-14}$$

We have

$$n(y) = \frac{(1, -\rho'(y_2))}{\sqrt{1 + (\rho'(y_2))^2}} \quad \text{for } y \in \Gamma_q \tag{10-15}$$

and

$$y_t = x_t - d_t n(y_t).$$

Thus

$$y_{t,2} = t + d_t \frac{\rho'(y_{t,2})}{\sqrt{1 + (\rho'(y_{t,2}))^2}}$$

and

$$y_{t,2} \le t + d_t < t + \eta t = (1+\eta)t.$$

Step 6. In this step we show that there exists $c_3 > 0$ and $t_3 > 0$ such that

$$\rho(y_{t,2})\rho'(y_{t,2}) \ge c_3 t^{2q-1} \quad \text{for } 0 < t < t_3. \tag{10-16}$$

Set $t_3 = \min(t_1, t_2)$. From (10-3), (10-8) and (10-12) it follows that

$$-n_2(y_t)y_{t,1} = \partial_{n(y_t)} v(y_t) \geq \frac{c_2}{d_t} v(x_t) \geq \frac{c_2}{d_t} c_1 t^{2q} \quad \text{for } 0 < t < t_3. \tag{10-17}$$

From (10-17), (10-15), (10-10) and (10-11) we get

$$\begin{aligned}\rho(y_{t,2})\rho'(y_{t,2}) = \rho'(y_{t,2})y_{t,1} &\geq \frac{\rho'(y_{t,2})}{\sqrt{1+(\rho'(y_{t,2}))^2}} y_{t,1} \\ &= -n_2(y_t)y_{t,1} \geq \frac{c_2}{d_t} c_1 t^{2q} \geq \frac{1}{\eta} c_1 c_2 t^{2q-1} = c_3 t^{2q-1}.\end{aligned}$$

Step 7. In this step, using the convexity of ρ we finish the proof of the lemma.

By the convexity of ρ, the function $\rho\rho'$ is nondecreasing; hence, by (10-14) and (10-16), we have

$$\rho((1+\eta)t)\rho'((1+\eta)t) \geq \rho(y_{t,2})\rho'(y_{t,2}) \geq c_3 t^{2q-1} \quad \text{for } 0 < t < t_3.$$

Letting $\tau = (1+\eta)t$ we have that

$$\rho(\tau)\rho'(\tau) \geq c_3 \Big(\frac{\tau}{1+\eta}\Big)^{2q-1} = c_4 \tau^{2q-1} \quad \text{for } 0 < \tau < (1+\eta)t_3 = \tau_0.$$

It follows that

$$(\rho^2)'(\tau) \geq 2c_4 \tau^{2q-1} \quad \text{for } 0 < \tau < \tau_0$$

and by integration we obtain

$$\rho(\tau) \geq c_5 \tau^q \quad \text{for } 0 < \tau < \tau_0.$$

From the convexity of ρ it follows that $\tau\rho'(\tau) \geq \rho(\tau)$; hence

$$\rho'(\tau) \geq c_5 \tau^{q-1} \quad \text{for } 0 < \tau < \tau_0$$

and this completes the proof of the lemma. □

Proof of Theorem 10. By Lemmas 47 and 48 we have that either $\rho = 0$ in $\big(0, \frac{1}{2}\delta\big)$ and $u = u_{hs}$ in $\Omega \cap B_\delta \cap \{x_1 > 0,\ x_2 > 0\}$ or, for all $q > 1$, there exist $c > 0$ and $t_0 > 0$ such that (10-2) holds.

In the latter case, if Γ is $C^{1,\alpha}$ regular for some $0 < \alpha < 1$ at the origin, then there exists $C > 0$ and $\delta_1 > 0$ such that

$$|\rho'(x_2) - \rho'(+0)| \leq C|x_2|^\alpha \quad \text{for } 0 < x_2 < \delta_1.$$

But, because $\rho'(+0) = 0$ and $\rho'(x_2) \geq 0$, we should have

$$\rho'(x_2) \leq C x_2^\alpha \quad \text{for } 0 < x_2 < \delta_1.$$

This contradicts with (10-2) if we take $1 < q < 1 + \alpha$. □

11. Further directions

The problem considered in this paper might be thought of as a prototype of free boundary problems, especially the obstacle problem, with a degenerate force term. There are many open questions in these problems and we are working to complete some works on these questions.

Some further directions are as follows:

(1) Higher dimension. It is interesting to consider the same problem in higher dimensions with possibly different dimensions for the set where the force term vanishes. In [Yeressian 2015] the key nondegeneracy result is proved for such higher-dimensional problems when the force term vanishes on a linear subspace.

(2) More general force terms. Partial results show that, when the force term is of the form $|x_1|^\alpha$ for $\alpha > 0$, the number of homogeneous global solutions — and together with it the possible Weiss balanced energy levels — grows linearly with $\alpha > 0$. Again in [Yeressian 2015] the key nondegeneracy result is proved for such general force terms. Many results in this paper could be written for such more general forces, but to have a reasonable bound on the size of the paper we have opted to consider the case $\alpha = 1$ only.

(3) Degenerate free boundary points and points where $W(+0, x, u) = 2W(1, u_{hs})$. We know that at these points the free boundary converges tangentially to the line $\{x_1 = 0\}$ and we know some topological structure of the set of these points based on the upper semicontinuity of the Weiss balanced energy. Also, in a particular case we have proved an irregularity result for the free boundary at such points. It is interesting to study the structure of the free boundary near to such points in more detail.

(4) Uniform results. For the nondegenerate obstacle problems there are many results which hold uniformly for a class of problems; see [Petrosyan et al. 2012]. But in this paper we have only considered a single solution alone.

(5) Parabolic problem. The problem considered in this paper has a parabolic analogue. It is interesting to know the exact influence of the degeneracy of the force term in the parabolic problems.

Acknowledgements

The author would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Free Boundary Problems and Related Topics, where work on this paper was undertaken. The author is grateful to Henrik Shahgholian for inspiring discussions and to Luis Caffarelli for his suggestion to prove the convergence of the normal of the free boundary to the normal of the free boundary of the blowup limit. The author would also like to thank the anonymous referees for their suggestions. The proof of the uniqueness of the blowup limits was significantly simplified and the version in this paper is due to one of them.

References

[Alt 1977] H. W. Alt, "The fluid flow through porous media: regularity of the free surface", *Manuscripta Math.* **21**:3 (1977), 255–272. MR 56 #7475 Zbl 0371.76080

[Caffarelli 1977] L. A. Caffarelli, "The regularity of free boundaries in higher dimensions", *Acta Math.* **139**:1 (1977), 155–184. MR 56 #12601 Zbl 0386.35046

[Caffarelli 1998] L. A. Caffarelli, "The obstacle problem revisited", *J. Fourier Anal. Appl.* **4**:4–5 (1998), 383–402. MR 2000b:49004 Zbl 0928.49030

[Han and Lin 2011] Q. Han and F. Lin, *Elliptic partial differential equations*, 2nd ed., Courant Lecture Notes in Mathematics **1**, Courant Institute of Mathematical Sciences, New York, 2011. MR 2012c:35077 Zbl 1210.35031

[Kinderlehrer and Stampacchia 1980] D. Kinderlehrer and G. Stampacchia, *An introduction to variational inequalities and their applications*, Pure and Applied Mathematics **88**, Academic Press, New York, 1980. MR 81g:49013 Zbl 0457.35001

[Petrosyan et al. 2012] A. Petrosyan, H. Shahgholian, and N. Uraltseva, *Regularity of free boundaries in obstacle-type problems*, Graduate Studies in Mathematics **136**, American Mathematical Society, Providence, RI, 2012. MR 2962060 Zbl 1254.35001

[Shahgholian and Uraltseva 2003] H. Shahgholian and N. Uraltseva, "Regularity properties of a free boundary near contact points with the fixed boundary", *Duke Math. J.* **116**:1 (2003), 1–34. MR 2003m:35253 Zbl 1050.35157

[Shahgholian et al. 2007] H. Shahgholian, N. Uraltseva, and G. S. Weiss, "The two-phase membrane problem—regularity of the free boundaries in higher dimensions", *Int. Math. Res. Not.* **2007**:8 (2007), art. ID rnm026. MR 2009b:35444 Zbl 1175.35157

[Varvaruca and Weiss 2011] E. Varvaruca and G. S. Weiss, "A geometric approach to generalized Stokes conjectures", *Acta Math.* **206**:2 (2011), 363–403. MR 2012g:35402 Zbl 1238.35194

[Varvaruca and Weiss 2012] E. Varvaruca and G. S. Weiss, "The Stokes conjecture for waves with vorticity", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **29**:6 (2012), 861–885. MR 2995099 Zbl 1317.35209

[Varvaruca and Weiss 2014] E. Varvaruca and G. S. Weiss, "Singularities of steady axisymmetric free surface flows with gravity", *Comm. Pure Appl. Math.* **67**:8 (2014), 1263–1306. MR 3225630 Zbl 1304.35524

[Weiss 1998] G. S. Weiss, "Partial regularity for weak solutions of an elliptic free boundary problem", *Comm. Partial Differential Equations* **23**:3–4 (1998), 439–455. MR 99d:35188 Zbl 0897.35017

[Weiss 1999] G. S. Weiss, "A homogeneity improvement approach to the obstacle problem", *Invent. Math.* **138**:1 (1999), 23–50. MR 2000h:35057 Zbl 0940.35102

[Yeressian 2015] K. Yeressian, "Nondegeneracy in the obstacle problem with a degenerate force term", *Interfaces Free Bound.* **17**:2 (2015), 233–244. MR 3391970 Zbl 1328.35337

Received 25 Oct 2014. Revised 27 Oct 2015. Accepted 6 Jan 2016.

KAREN YERESSIAN: karen.yeressian@math.uzh.ch
Institute of Mathematics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.439

A COUNTEREXAMPLE TO THE HOPF–OLEINIK LEMMA (ELLIPTIC CASE)

Darya E. Apushkinskaya and Alexander I. Nazarov

Dedicated to Professor M.V. Safonov

We construct a new counterexample to the Hopf–Oleinik boundary point lemma. It shows that for convex domains, the $C^{1,\mathrm{Dini}}$ assumption on $\partial\Omega$ is the necessary and sufficient condition providing the estimates of Hopf–Oleinik type.

1. Introduction

The influence of the properties of a domain on the behavior of a solution is one of the most important topics in the qualitative analysis of partial differential equations.

The significant result in this field is the Hopf–Oleinik lemma, known also as the "boundary point principle". This celebrated lemma states:

Lemma. *Let u be a nonconstant solution to a second-order homogeneous uniformly elliptic nondivergence equation with bounded measurable coefficients, and let u attend its extremum at a point x^0 located on the boundary of a domain $\Omega \subset \mathbb{R}^n$. Then $(\partial u/\partial \boldsymbol{n})(x^0)$ is necessarily nonzero provided that $\partial\Omega$ satisfies the proper assumptions at x^0.*

This result was established in a pioneering paper of S. Zaremba [1910] for the Laplace equation in a 3-dimensional domain Ω having an interior touching ball at x^0 and generalized by G. Giraud [1932; 1933] to equations with Hölder-continuous leading coefficients and continuous lower-order coefficients in domains Ω belonging to the class $C^{1,\alpha}$ with $\alpha \in (0, 1)$.

Notice that a related assertion about the negativity on $\partial\Omega$ of the normal derivative of the Green's function corresponding to the Dirichlet problem for the Laplace operator was proved much earlier for 2-dimensional smooth domains by C. Neumann [1888] (see also [Korn 1901]). The result of [Neumann 1888] was extended for operators with lower-order coefficients by L. Lichtenstein [1924]. The same version of the boundary point principle for the Laplacian and 3-dimensional domains satisfying a more flexible interior paraboloid condition was obtained by M. V. Keldysch and M. A. Lavrent'ev [1937].

A crucial step in studying the boundary point principle was made by E. Hopf [1952] and O. A. Oleĭnik [1952], who simultaneously and independently proved the statement for the general elliptic equations with bounded coefficients and domains satisfying an interior ball condition at x^0.

Later the efforts of many mathematicians were focused on the generalization of the boundary point principle in several directions (for the details, we refer the reader to [Alvarado et al. 2011; Alvarado

MSC2010: 35J15, 35B45.
Keywords: elliptic equations, Hopf–Oleinik lemma, Dini continuity, counterexample.

2011] and references therein). Among these directions are the extension of the class of operators and the class of solutions, as well as the weakening of assumptions on the boundary.

The widening of the class of operators to singular/degenerate ones was made in the papers [Kamynin and Himčenko 1975; 1977; Alvarado et al. 2011], while the uniform elliptic operators with unbounded lower-order coefficients were studied in [Safonov 2010; Nazarov 2012] (see also [Nazarov and Uraltseva 2009]). We mention also the publications [Tolksdorf 1983; Mikayelyan and Shahgholian 2015], where the boundary point principle was established for a class of degenerate quasilinear operators including the p-Laplacian.

We note that before 2010, all the results were formulated for classical solutions, i.e., $u \in C^2(\Omega)$. The class of solutions was expanded in [Safonov 2010] to strong generalized solutions with Sobolev's second-order derivatives. The latter requirement seems to be natural in the study of nondivergent elliptic equations.

The reduction of the assumptions on the boundary of Ω up to $C^{1,\mathrm{Dini}}$-regularity was realized for various elliptic operators in the papers [Widman 1967; Himčenko 1970; Lieberman 1985] (see also [Safonov 2008]). A weakened form of the Hopf–Oleinik lemma (the existence of a boundary point x^1 in any neighborhood of x^0 and a direction ℓ such that $(\partial u/\partial \ell)(x^1) \neq 0$) was proved in [Nadirashvili 1983] for a much wider class of domains including all Lipschitz ones. We mention also the paper [Sweers 1997], where the behavior of superharmonic functions near the boundary of a 2-dimensional domain with corners is described in terms of the main eigenfunction of the Dirichlet Laplacian.

The sharpness of some requirements was confirmed by corresponding counterexamples constructed in [Widman 1967; Himčenko 1970; Kamynin and Himčenko 1975; Safonov 2008; Alvarado et al. 2011; Nazarov 2012]. In particular, the counterexamples from [Widman 1967; Himčenko 1970; Safonov 2008] show that the Hopf–Oleinik result fails for domains lying entirely in non-Dini paraboloids.

The main result of our paper is a new counterexample (see Theorem 4.2) showing the sharpness of the Dini condition for the boundary of Ω. The simplest version of this counterexample can be formulated as follows:

Counterexample. *Let Ω be a convex domain in $\mathbb{R}^n$, let $\partial\Omega$ in a neighborhood of the origin be described by the equation $x_n = F(x')$ with $F \geqslant 0$ and $F(0) = 0$, and let $u \in W^2_{n,\mathrm{loc}}(\Omega) \cap C(\overline{\Omega})$ be a solution of the uniformly elliptic equation*

$$-a^{ij}(x) D_i D_j u = 0 \quad \text{in } \Omega.$$

Suppose also that $u|_{\partial\Omega}$ vanishes at a neighborhood of the origin. If, in addition, the function

$$\delta(r) = \sup_{|x'| \leqslant r} \frac{F(x')}{|x'|}$$

is not Dini-continuous at zero, then $(\partial u/\partial \boldsymbol{n})(0) = 0$.

Thus, it turns out that for convex domains, the Dini-continuity assumption on $\delta(r)$ is necessary and sufficient for the validity of the boundary point principle. We emphasize that in our counterexample the Dini condition fails for the supremum of $F(x')/|x'|$, while in all the previous results of this kind, it fails for the infimum of $F(x')/|x'|$. In other words, we show that violating the Dini condition just in one direction causes the failure of the Hopf–Oleinik lemma.

Notation and conventions. Throughout the paper we use the following notation:

- $x = (x', x_n) = (x_1, \dots, x_{n-1}, x_n)$ is a point in $\mathbb{R}^n$.
- $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_n > 0\}$.
- $|x|, |x'|$ are the Euclidean norms in the corresponding spaces.
- $\chi_{\mathcal{E}}$ denotes the characteristic function of the set $\mathcal{E} \subset \mathbb{R}^n$.
- Ω is a bounded domain in $\mathbb{R}^n$ with boundary $\partial\Omega$.
- $\mathcal{P}_{r,h}(\bar{x}') = \{x \in \mathbb{R}^n : |x' - \bar{x}'| < r,\ 0 < x_n < h\}$ and $\mathcal{P}_r(\bar{x}') = \mathcal{P}_{r,r}(\bar{x}')$.
- $\mathcal{P}_{r,h} = \mathcal{P}_{r,h}(0)$ and $\mathcal{P}_r = \mathcal{P}_r(0)$.
- $B_r(x^0)$ is the open ball in $\mathbb{R}^n$ with center x^0 and radius r; $B_r = B_r(0)$.
- For $r_1 < r_2$, we define the annulus $\mathcal{B}(x^0, r_1, r_2) = B_{r_2}(x^0) \setminus \overline{B_{r_1}(x^0)}$.
- $v_+ = \max\{v, 0\}$ and $v_- = \max\{-v, 0\}$.
- $\|\cdot\|_{\infty,\Omega}$ denotes the norm in $L_\infty(\Omega)$.
- We adopt the convention that the indices i and j run from 1 to n. We also adopt the convention regarding summation with respect to repeated indices.
- D_i denotes the operator of (weak) differentiation with respect to x_i.
- $D = (D', D_n) = (D_1, \dots, D_{n-1}, D_n)$.
- $\mathcal{L}$ is a linear uniformly elliptic operator with measurable coefficients

$$\mathcal{L}u \equiv -a^{ij}(x) D_i D_j u + b^i(x) D_i u, \quad \nu \mathcal{I}_n \le (a^{ij}(x)) \le \nu^{-1} \mathcal{I}_n, \tag{1}$$

  where $\mathcal{I}_n$ is the $n \times n$ identity matrix. We define $\boldsymbol{b}(x) = (b^1(x), \dots, b^n(x))$.
- We use the letters C and N (with or without indices) to denote various constants. To indicate that, say, C depends on some parameters, we list them in parentheses: $C(\cdots)$.

Definition 1.1. We say that a function $\sigma : [0, 1] \to \mathbb{R}_+$ belongs to the class $\mathcal{D}_1$ if

- σ is increasing, $\sigma(0) = 0$, and $\sigma(1) = 1$;
- $\sigma(t)/t$ is summable and decreasing.

Remark 1.2. Our assumption about the decay of $\sigma(t)/t$ is not restrictive. Indeed, for any increasing function $\sigma : [0, 1] \to \mathbb{R}_+$ satisfying $\sigma(0) = 0$ and $\sigma(1) = 1$ and having summable $\sigma(t)/t$, we can define

$$\tilde{\sigma}(t) = t \sup_{\tau \in [t,1]} \frac{\sigma(\tau)}{\tau}, \quad t \in (0, 1).$$

It is easy to see that $\tilde{\sigma} \in \mathcal{D}_1$, $\tilde{\sigma}(t)/t$ decreases and $\sigma(t) \leqslant \tilde{\sigma}(t)$ for all $t \in (0, 1]$.

Definition 1.3. Let a function σ belong to the class $\mathcal{D}_1$. We define the function $\mathcal{J}_\sigma$ as

$$\mathcal{J}_\sigma(s) := \int_0^s \frac{\sigma(\tau)}{\tau}\, d\tau. \tag{2}$$

Remark 1.4. The decreasing of $\sigma(t)/t$ implies

$$\sigma(t) \leqslant \mathcal{J}_\sigma(t) \qquad \forall t \in [0,1]. \tag{3}$$

In addition, for $t \leqslant t_0 \leqslant 1$, we have

$$\sigma(t/t_0) = \frac{\sigma(t/t_0)}{t/t_0} \cdot t/t_0 \leqslant \frac{\sigma(t)}{t} \cdot t/t_0 = \frac{\sigma(t)}{t_0}, \tag{4}$$

and, similarly,

$$\mathcal{J}_\sigma(t/t_0) \leqslant \frac{\mathcal{J}_\sigma(t)}{t_0}. \tag{5}$$

Definition 1.5. We say that a function ζ satisfies the Dini condition at zero if

$$|\zeta(r)| \leqslant C\sigma(r),$$

and σ belongs to the class $\mathcal{D}_1$.

2. Preliminaries

Properties of Ω. Let Ω be a bounded domain in $\mathbb{R}^n$. Without loss of generality, we may assume $0 \in \partial\Omega$.

Suppose that Ω is locally convex in a neighborhood of the origin. Without restriction, the latter means that for some $0 < \mathcal{R}_0 \leqslant 1$, we have

$$\mathcal{P}_{\mathcal{R}_0} \cap \Omega = \big\{(x', x_n) \in \mathbb{R}^n \,:\, |x'| \leqslant \mathcal{R}_0,\ F(x') < x_n < \mathcal{R}_0\big\},$$

where F is a convex nonnegative function satisfying $F(0) = 0$.

For $r \in (0, \mathcal{R}_0)$, we define the functions $\delta = \delta(r)$ and $\delta_1 = \delta_1(r)$ by the formulas

$$\delta(r) := \max_{|x'| \leqslant r} \frac{F(x')}{|x'|}, \quad \delta_1(r) := \max_{|x'| \leqslant r} |\nabla F(x')|. \tag{6}$$

Lemma 2.1. *The following statements hold*:

(a) $\delta_1(r) \to 0$ *as* $r \to 0$ *if and only if* $\delta(r) \to 0$ *as* $r \to 0$.

(b) $\delta_1(r)$ *satisfies the Dini condition at zero if and only if* $\delta(r)$ *satisfies the Dini condition at zero.*

Proof. By the convexity of F, we have for any x' and z', the estimate

$$F(z') \geqslant F(x') + \nabla F(x') \cdot (z' - x'). \tag{7}$$

Therefore,

$$|\nabla F(x')| \geqslant \nabla F(x') \cdot \frac{x'}{|x'|} \geqslant \frac{F(x')}{|x'|},$$

and, consequently,

$$\delta_1(r) \geqslant \delta(r). \tag{8}$$

On the other hand, for any $r < \frac{1}{2}\mathcal{R}_0$, we can find a point x'_* such that

$$|\nabla F(x'_*)| = \delta_1(r).$$

Choosing

$$z' = x'_* + r\,\frac{\nabla F(x'_*)}{|\nabla F(x'_*)|},$$

we easily deduce from (7) the inequalities

$$|z'| \leqslant 2r \quad \text{and} \quad F(z') \geqslant r\delta_1(r),$$

which provide

$$\delta(2r) \geqslant \delta(|z'|) \geqslant \tfrac{1}{2}\delta_1(r). \tag{9}$$

Combining (8) and (9), we conclude that statement (a) is obvious and the integrals

$$\int_0^{\mathcal{R}_0} \frac{\delta(r)}{r}\,dr \quad \text{and} \quad \int_0^{\mathcal{R}_0} \frac{\delta_1(r)}{r}\,dr$$

converge simultaneously. □

If $\delta(r)$ does not converge to zero as $r \to 0$, we can easily see that the domain Ω is contained in a dihedral wedge with the angle less than π and the edge going through the origin. For this case, the statement of Theorem 4.2 is proved already in [Apushkinskaya and Nazarov 2000, Theorem 4.3]. For this reason, we will assume throughout this paper that

$$\delta(r) \to 0 \quad \text{as } r \to 0. \tag{10}$$

In view of (10), it is evident that δ and δ_1 are moduli of continuity at the origin of the functions $F(x')/|x'|$ and $|\nabla F(x')|$, respectively.

Properties of $\mathcal{X}(\Omega)$. Let $\mathcal{X}(\Omega)$ be a function space with the norm $\|\cdot\|_{\mathcal{X},\Omega}$. For $\Omega_1 \subset \Omega$, we will assume

$$\|f\|_{\mathcal{X},\Omega_1} = \|f \cdot \chi_{\Omega_1}\|_{\mathcal{X},\Omega}.$$

We suppose that $\mathcal{X}(\Omega)$ has the following properties:

(i) For an arbitrary measurable function g defined in Ω and any function $f \in \mathcal{X}(\Omega)$, the inequality $|g(x)| \leqslant |f(x)|$ implies $g \in \mathcal{X}(\Omega)$ and $\|g\|_{\mathcal{X},\Omega} \leqslant \|f\|_{\mathcal{X},\Omega}$.

(ii) For $f_k \in \mathcal{X}(\Omega)$, the convergence $f_k \searrow 0$ a.e. in Ω implies $\|f_k\|_{\mathcal{X},\Omega} \to 0$.

Using the terminology of the classic monograph of Kantorovich and Akilov [1982], we may say that $\mathcal{X}(\Omega)$ is the ideal functional space with order continuous monotone norm (see [Kantorovich and Akilov 1982, §3, Chapter IV, Part I] for more details).

We will also assume that

(iii) $\mathcal{X}_{\text{loc}}(\Omega)$ contains the Orlicz space $L_{\Phi,\text{loc}}(\Omega)$ with $\Phi(\xi) = e^{\xi} - \xi - 1$.

Finally, the basic assumption about $\mathcal{X}(\Omega)$ is the Aleksandrov-type maximum principle. Namely, we denote by $\mathcal{W}^2_{\mathcal{X},\text{loc}}(\Omega)$ the set of the functions u satisfying $D(Du) \in \mathcal{X}_{\text{loc}}(\Omega)$, and suppose that if $u \in \mathcal{W}^2_{\mathcal{X},\text{loc}}(\Omega) \cap \mathcal{C}(\overline{\Omega})$, $u|_{\partial\Omega} \leq 0$, and $|\boldsymbol{b}| \in \mathcal{X}(\Omega)$ then

$$u \leqslant N_0(n, \nu, \|\boldsymbol{b}\|_{\mathcal{X},\Omega}) \cdot \operatorname{diam}(\Omega) \cdot \|(\mathcal{L}u)_+\|_{\mathcal{X},\{u>0\}}. \tag{11}$$

Remark 2.2. It is well known from [Aleksandrov 1960; 1963; Bakel′man 1961] (see also the survey [Nazarov 2005] for further references) that $L_n(\Omega)$ has property (11). It is also evident that properties (i)–(iii) are satisfied in $L_n(\Omega)$. Therefore, $L_n(\Omega)$ can be treated as a "basic" example of $\mathcal{X}(\Omega)$. As other examples of the space $\mathcal{X}(\Omega)$, we mention some Lebesgue weighted spaces with power weights (see [Nazarov 2001]).

Remark 2.3. Unlike the natural properties (i)–(ii), assumption (iii) is a rather "technical" one. Without (iii), our arguments from the proof of Step 3 in Theorem 4.1 are not applicable to the approximating operator $\mathcal{L}_\varepsilon$. So, we cannot withdraw (iii) in abstract setting. However, in all known examples of $\mathcal{X}(\Omega)$, property (iii) is satisfied.

Remark 2.4. Some of the statements that will be referred to in the sequel were proved earlier just for the case $\mathcal{X}(\Omega) = L_n(\Omega)$. However, if all the arguments are based only on the Aleksandrov-type maximum principle, these statements remain valid for an arbitrary considered space $\mathcal{X}(\Omega)$. In such cases, we will refer to this remark without any further explanation.

We also need the following convergence lemmas.

Lemma 2.5. *Let $\{f_j\}$ be a sequence of measurable functions on Ω, and let $f \in \mathcal{X}(\Omega)$. Suppose also that $f_j \to 0$ in measure on Ω, and $|f_j(x)| \leqslant |f(x)|$.*

Then

$$\|f_j\|_{\mathcal{X},\Omega} \to 0 \quad \text{as } j \to \infty. \tag{12}$$

Proof. We argue by a contradiction. Suppose (12) fails. Then there exists a subsequence $\{f_{j_k}\}$ satisfying

$$\|f_{j_k}\|_{\mathcal{X},\Omega} \geqslant \varepsilon > 0 \quad \forall k \in \mathbb{N}. \tag{13}$$

Due to the Riesz theorem, there exists also a subsubsequence $\{f_{j_{k_l}}\}$ such that

$$f_{j_{k_l}} \to 0 \quad \text{a.e. in } \Omega.$$

For simplicity of notation, we renumber the latter subsequence $\{f_{j_{k_l}}\}$ and denote its elements again by f_j.

Setting $\tilde{f}_k := \sup_{j\geqslant k}|f_j|$, we can easily see that $\tilde{f}_k \searrow 0$ a.e. in Ω. Now, taking into account properties (i) and (ii) of the space $\mathcal{X}(\Omega)$, we immediately get a contradiction with inequalities (13). □

Lemma 2.6. *Let $f \in \mathcal{X}(\Omega)$, and let $\mu(\rho) := \sup\limits_{x\in\Omega} \|f\|_{\mathcal{X},B_\rho(x)\cap\Omega}$.*

Then

$$\mu(\rho) \to 0 \quad \text{as } \rho \to 0.$$

Proof. For every $\rho > 0$, there exists a point $x^* = x^*(\rho) \in \Omega$ such that

$$\|f\|_{\mathcal{X},B_\rho(x^*)\cap\Omega} \geqslant \tfrac{1}{2}\mu(\rho).$$

Next, for the sequence $f_\rho := f \cdot \chi_{B_\rho(x^*)}$, it is evident that $|f_\rho| \to 0$ in measure on Ω. An application of Lemma 2.5 finishes the proof. □

Remark 2.7. We call $\mu(\rho) := \sup_{x\in\Omega} \|f\|_{\mathcal{X} B_\rho(x)\cap\Omega}$ the modulus of continuity of the function f in $\mathcal{X}(\Omega)$.

Lemma 2.8. *Let $D(Du) \in \mathcal{X}(\Omega)$, let $\mathcal{L}$ be defined by* (1), *and let $\mathcal{L}u \in \mathcal{X}(\Omega)$. There exists the family of operators*

$$\mathcal{L}_\varepsilon = -a_\varepsilon^{ij}(x) D_i D_j + b_\varepsilon^i(x) D_i$$

with smooth coefficients a_ε^{ij} and bounded coefficients b_ε^i satisfying

$$\nu \mathcal{I}_n \le (a_\varepsilon^{ij}(x)) \le \nu^{-1}\mathcal{I}_n, \quad x \in \Omega, \tag{14}$$

$$|b_\varepsilon^i(x)| \leqslant |b^i(x)|, \quad x \in \Omega, \tag{15}$$

$$\|(\mathcal{L}-\mathcal{L}_\varepsilon)u\|_{\mathcal{X},\Omega} \to 0 \quad \text{as } \varepsilon \to 0. \tag{16}$$

Proof. We start with the extension of a^{ij} on the whole $\mathbb{R}^n$ by the identity matrix and denote by a_ε^{ij} the standard mollification of extended functions a^{ij}. By construction, the coefficients a_ε^{ij} are smooth functions converging as $\varepsilon \to 0$ to a^{ij} a.e. in Ω. Moreover, it is clear that inequalities (14) are true.

Further, we set

$$\tilde{b}_\varepsilon^i(x) := \min\{|b^i(x)|, \varepsilon^{-1}\} \cdot \operatorname{sign} b^i(x). \tag{17}$$

In view of (17), it is evident that $\tilde{b}_\varepsilon^i D_i u$ converges as $\varepsilon \to 0$ to $b^i D_i u$ a.e. in Ω. We claim that it is possible to change $\tilde{b}_\varepsilon^i$ such that the "corrected coefficients" b_ε^i satisfy

$$|b_\varepsilon^i D_i u| \leqslant |b^i D_i u| \quad \text{in } \Omega. \tag{18}$$

Indeed, if $|\tilde{b}_\varepsilon^i D_i u| \leqslant |b^i D_i u|$ in Ω then (18) holds with $b_\varepsilon^i \equiv \tilde{b}_\varepsilon^i$. Otherwise, consider a point $x^0 \in \Omega$, where $|\tilde{b}_\varepsilon^i(x^0) D_i u(x^0)| > |b^i(x^0) D_i u(x^0)|$.

(a) Let $\tilde{b}_\varepsilon^i(x^0) D_i u(x^0) > b^i(x^0) D_i u(x^0) \geqslant 0$. In this case, we decrease all the coefficients $\tilde{b}_\varepsilon^i(x^0)$ corresponding to the positive summands such that the sums $b_\varepsilon^i D_i u$ and $b^i D_i u$ become equal.

(b) Let $\tilde{b}_\varepsilon^i(x^0) D_i u(x^0) < b^i(x^0) D_i u(x^0) \leqslant 0$. In this case, we decrease all the coefficients $\tilde{b}_\varepsilon^i(x^0)$ corresponding to the negative summands such that the sums $b_\varepsilon^i D_i u$ and $b^i D_i u$ become equal.

(c) Finally, let $\tilde{b}_\varepsilon^i(x^0) D_i u(x^0)$ and $b^i(x^0) D_i u(x^0)$ have different signs. In this case, we apply to $-b_\varepsilon^i(x^0)$ the arguments from case (a) or from case (b), respectively.

Due to construction, the "corrected sum" $b_\varepsilon^i D_i u$ also converges as $\varepsilon \to 0$ to $b^i D_i u$ a.e. in Ω, and the pointwise inequalities (15) hold true.

Finally, taking into account (18) and applying Lemma 2.5, we get (16). ☐

3. Gradient estimates near the boundary

Lemma 3.1. *Let $\mathcal{N} \subset \mathbb{R}_+^n$ be an open set, let $\gamma = \nu/\sqrt{n-1}$, let $\rho > 0$, and let*

$$\Pi_\rho = \{y \in \mathbb{R}^n : |y_i| < \rho \text{ for } i = 1, \dots, n-1;\ 0 < y_n < \gamma\rho\}.$$

We assume that $|\boldsymbol{b}| \in \mathcal{X}(\mathcal{N})$ and a function v satisfies the conditions

$$v \in \mathcal{W}^2_{\mathcal{X},\mathrm{loc}}(\mathcal{N}), \qquad v \geqslant 0 \quad \text{in } \Pi_\rho, \qquad v \geqslant k = \text{constant} > 0 \quad \text{on } \partial\mathcal{N} \cap \overline{\Pi}_\rho.$$

Then

$$v \geqslant C_1 k - C_2 k \|\boldsymbol{b}\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho} - C_3 \rho \|(\mathcal{L}v)_-\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho} \quad \text{in } \mathcal{N}\cap B_{\frac{\gamma\rho}{4}}(z),$$

where $z=(0,\dots,0,\frac{1}{2}\gamma\rho)$, *while* $C_1=\frac{1}{16}(1-\gamma^2)$, $C_2=C_2(n,\nu,\|\boldsymbol{b}\|_{\mathcal{X},\mathcal{N}})$, *and* $C_3 = C_3(n,\nu,\|\boldsymbol{b}\|_{\mathcal{X},\mathcal{N}})$.

Proof. The proof is similar in spirit to [Apushkinskaya and Ural′tseva 1995, Lemma 1].

Consider the barrier function

$$\psi(y)=k\left(\left(1-\frac{y_n}{\gamma\rho}\right)^2-\frac{|y'|^2}{\rho^2}\right).$$

An elementary computation gives

$$\mathcal{L}\psi \leqslant k\left(\frac{2(n-1)}{\rho^2}\nu^{-1}-\frac{2}{\gamma^2\rho^2}\nu\right)+|\boldsymbol{b}||D\psi| \leqslant N_1(n,\nu)|\boldsymbol{b}|\frac{k}{\rho} \quad \text{in } \Pi_\rho.$$

Moreover, setting

$$\mathcal{S}_1=\{y\in\partial(\mathcal{N}\cap\Pi_\rho) : |y_i|=\rho \text{ for some } i=1,\dots,n-1\},$$
$$\mathcal{S}_2=\{y\in\partial(\mathcal{N}\cap\Pi_\rho) : y_n=\gamma\rho\},$$

we have

$$\psi\big|_{\mathcal{S}_1\cup\mathcal{S}_2}\leqslant 0\leqslant v,$$
$$\psi\big|_{\partial\mathcal{N}\cap\overline{\Pi}_\rho}\leqslant k\leqslant v\big|_{\partial\mathcal{N}\cap\overline{\Pi}_\rho}.$$

Applying inequality (11) in $\mathcal{N}\cap\Pi_\rho$ to the difference $\psi-v$, we obtain

$$\psi-v\leqslant N_0\cdot\operatorname{diam}(\Pi_\rho)\cdot\|(\mathcal{L}\psi-\mathcal{L}v)_+\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho} \quad \text{in } \mathcal{N}\cap\Pi_\rho,$$

and, consequently,

$$\begin{aligned} v &\geqslant k\left(\left(1-\frac{\frac{3}{4}\gamma\rho}{\gamma\rho}\right)^2-\frac{\gamma^2\rho^2}{16\rho^2}\right)-C_2k\|\boldsymbol{b}\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho}-C_3\rho\|(\mathcal{L}v)_-\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho} \\ &= \tfrac{1}{16}(1-\gamma^2)k-C_2k\|\boldsymbol{b}\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho}-C_3\rho\|(\mathcal{L}v)_-\|_{\mathcal{X},\mathcal{N}\cap\Pi_\rho} \quad \text{in } \mathcal{N}\cap B_{\frac{\gamma\rho}{4}}(z). \end{aligned}$$ ☐

Our next statement is a version of [Nazarov 2012, Theorem 2.3].

Lemma 3.2. *Let* $v\in\mathcal{W}^2_{\mathcal{X},\mathrm{loc}}(\Omega)\cap\mathcal{C}(\overline{\Omega})$, *let* $v|_{\partial\Omega}=0$, *and let* $|\boldsymbol{b}|\in\mathcal{X}(\Omega)$. *Suppose also that for all* $\rho\leqslant\rho_*\leqslant 1$, *the inequalities*

$$\|b^n\|_{\mathcal{X},\mathcal{P}_\rho\cap\Omega}\leqslant\mathfrak{B}\sigma(\rho/\rho_*), \quad \|(\mathcal{L}v)_+\|_{\mathcal{X},\mathcal{P}_\rho\cap\Omega}\leqslant\mathfrak{F}\sigma(\rho/\rho_*)$$

hold true. Here $\mathfrak{B}$ *and* $\mathfrak{F}$ *are some positive constants, while the function* σ *belongs to* $\mathcal{D}_1$.

Then

$$\sup_{0<x_n<\rho}\frac{v(0,x_n)}{x_n}\leqslant C_4\big(\rho^{-1}\sup_{\mathcal{P}_\rho\cap\Omega}v+\mathfrak{F}\mathcal{J}_\sigma(\rho/\rho_*)\big) \quad \forall\rho\leqslant\rho_*. \tag{19}$$

Here the constant C_4 *depends on* n, ν, $\mathfrak{B}$, σ, *and on the moduli of continuity of* $|\boldsymbol{b}'|$ *in* $\mathcal{X}(\mathcal{P}_{\rho_*}\cap\Omega)$, *whereas* $\mathcal{J}_\sigma$ *is a function defined by formula* (2).

Remark 3.3. We recall that $0 \in \partial\Omega$.

Proof. First, we assume that $\rho \leqslant \bar\rho$, where $\bar\rho \leqslant \rho_*$ will be fixed later. Following [Nazarov 2012], we introduce the sequence of cylinders $\mathcal{P}_{\rho_k,h_k}$, with $k \geqslant 0$, where $\rho_k = 2^{-k}\rho$ and $h_k = \zeta_k\rho_k$, while the sequence $\zeta_k \downarrow 0$ will be chosen later.

We set $w_k = v - M_k x_n$, where the quantities M_k, with $k \geqslant 1$, are defined as

$$M_k = \sup_{\mathcal{P}_{\rho_k,h_{k-1}}\cap\Omega} \frac{v(x)}{\max\{x_n, h_k\}} \geqslant \sup_{\{\mathcal{P}_{\rho_k,h_{k-1}}\setminus\mathcal{P}_{\rho_k,h_k}\}\cap\Omega} \frac{v(x)}{x_n}.$$

It is easy to see that $w_k \leqslant 0$ on $\partial\Omega \cap \overline{\mathcal{P}}_{\rho_k,h_k}$, while the definition of M_k gives $w_k \leqslant 0$ on the top of the cylinder $\mathcal{P}_{\rho_k,h_k}$.

Let $x^0 \in \mathcal{P}_{\rho_k-h_k,h_k} \cap \Omega$. Taking into account Remark 2.4, we apply the so-called "boundary growth lemma" (see, for instance, [Ladyzhenskaya and Ural′tseva 1985, Lemma 2.5′], [Safonov 2010, Lemma 2.6] or [Nazarov 2012, Lemma 2.2]) to the (positive) function $M_k h_k - w_k$ in $\mathcal{P}_{h_k}(x^{0\prime}) \cap \Omega$. It gives for $x \in \mathcal{P}_{h_k/2,h_k}(x^{0\prime}) \cap \Omega$,

$$M_k h_k - w_k(x) \geqslant M_k h_k\big[\vartheta - N_2\|\boldsymbol{b}\|_{\mathcal{X},\mathcal{P}_{\rho_k}\cap\Omega}\big] - N_3 h_k\|(\mathcal{L}w_k)_+\|_{\mathcal{X},\mathcal{P}_{h_k}(x^{0\prime})\cap\Omega}, \tag{20}$$

where $\vartheta = \vartheta(n,\nu,\sigma,\mathfrak{B}) \in (0,1)$ and the positive constant N_2 depends on the same parameters as ϑ, whereas the positive constant N_3 is completely defined by the values of n, ν and $\mathfrak{B}$. We suppose that $\bar\rho$ is so small that the quantity in the square brackets is greater than $\vartheta/2$. Further, direct calculation shows that the assumptions of our lemma imply

$$\begin{aligned}\|(\mathcal{L}w_k)_+\|_{\mathcal{X},\mathcal{P}_{h_k}(x^{0\prime})\cap\Omega} &\leqslant \|(\mathcal{L}v)_+\|_{\mathcal{X},\mathcal{P}_{h_k}(x^{0\prime})\cap\Omega} + M_k\|b^n\|_{\mathcal{X},\mathcal{P}_{h_k}(x^{0\prime})\cap\Omega}\\ &\leqslant (\mathfrak{F} + M_k\mathfrak{B})\sigma(\rho_k/\rho_*).\end{aligned}$$

Substituting the last inequality into (20) and taking the supremum with respect to x^0, we obtain

$$\sup_{\mathcal{P}_{\rho_k-h_k,h_k}\cap\Omega} w_k \leqslant M_k h_k\big(1 - \vartheta/2 + N_2\mathfrak{B}\sigma(\rho_k/\rho_*)\big) + N_3 h_k\mathfrak{F}\sigma(\rho_k/\rho_*).$$

Repeating previous arguments provides for all integers $m \leqslant \rho_k/h_k$ the inequalities

$$\sup_{\mathcal{P}_{\rho_k-mh_k,h_k}\cap\Omega} w_k \leqslant M_k h_k\left((1-\vartheta/2)^m + N_2\mathfrak{B}\,\frac{\sigma(\rho_k/\rho_*)}{\vartheta/2}\right) + N_3 h_k\mathfrak{F}\,\frac{\sigma(\rho_k/\rho_*)}{\vartheta/2}.$$

Setting $m = \lfloor \rho_{k+1}/h_k\rfloor$, we arrive at

$$\sup_{\mathcal{P}_{\rho_{k+1},h_k}\cap\Omega} w_k \leqslant \frac{M_k h_k}{1-\vartheta/2}\left(\exp\left(-\lambda\frac{\rho_{k+1}}{h_k}\right) + N_2\mathfrak{B}\,\frac{\sigma(\rho_k/\rho_*)}{\vartheta/2}\right) + N_3 h_k\mathfrak{F}\,\frac{\sigma(\rho_k/\rho_*)}{(1-\vartheta/2)\vartheta/2},$$

where $\lambda = -\ln(1-\vartheta/2) > 0$.

Therefore, for $x \in \mathcal{P}_{\rho_{k+1},h_k} \cap \Omega$,

$$\frac{w_k(x)}{\max\{x_n, h_{k+1}\}} \leqslant M_k\gamma_k + N_3\mathfrak{F}\,\frac{\sigma(\rho_k/\rho_*)}{(1-\vartheta/2)\vartheta/2}\cdot\frac{2\zeta_k}{\zeta_{k+1}}, \tag{21}$$

where

$$\gamma_k = \frac{1}{1-\vartheta/2}\frac{2\zeta_k}{\zeta_{k+1}} \cdot \left(\exp\left(-\frac{\lambda}{2\zeta_k}\right) + N_2\mathfrak{B}\,\frac{\sigma(\rho_k/\rho_*)}{\vartheta/2}\right).$$

Estimate (21) implies

$$\begin{aligned} M_{k+1} &\leqslant M_k(1+\gamma_k) + N_3\mathfrak{F}\,\frac{\sigma(\rho_k/\rho_*)}{(1-\vartheta/2)\vartheta/2}\cdot\frac{2\zeta_k}{\zeta_{k+1}} \\ &\leqslant M_1 \cdot \prod_{j=1}^{k}(1+\gamma_j) + 2N_3\mathfrak{F}\cdot\sum_{j=1}^{k}\sigma(\rho_j/\rho_*)\frac{\zeta_j}{\zeta_{j+1}}\cdot\prod_{l=j}^{k-1}(1+\gamma_l). \end{aligned}$$

We set $\zeta_k = 1/(k+k_0)$ and choose k_0 so large and $\bar{\rho}/\rho_*$ so small that $\gamma_1 \leqslant \frac{1}{2}$. Note that $k_0 = k_0(n,\nu,\sigma,\mathfrak{B})$, whereas $\bar{\rho}/\rho_*$ depends on the same parameters as k_0 and, in addition, on the moduli of continuity of $|\boldsymbol{b}'|$ in $\mathcal{X}(\mathcal{P}_{\rho_*}\cap\Omega)$.

Now we observe that the first term in γ_k forms a convergent series. The same is true for the second term, since

$$\sum_{k=1}^{\infty}\sigma(2^{-k}\rho/\rho_*) \asymp \int_0^{\infty}\sigma(2^{-s}\rho/\rho_*)\,ds \asymp \mathcal{J}_\sigma(\rho/\rho_*).$$

Therefore, the infinite product $\Pi = \prod_k(1+\gamma_k)$ also converges, and we obtain for $k>1$, the inequality

$$\begin{aligned} M_k &\leqslant \Pi\cdot\left(M_1 + 2N_3\mathfrak{F}\cdot\sum_{j=1}^{k}\sigma(\rho_j/\rho_*)\frac{\zeta_j}{\zeta_{j+1}}\right) \\ &\leqslant \Pi\cdot\big(M_1 + N_4(n,\nu,\sigma,\mathfrak{B})\mathfrak{F}\,\mathcal{J}_\sigma(\rho/\rho_*)\big). \end{aligned} \tag{22}$$

Thus, all M_k are bounded. It remains only to note that

$$M_1 \leqslant \frac{1}{h_1}\sup_{\mathcal{P}_{\rho/2}\cap\Omega} v. \tag{23}$$

Combining (22) and (23), we arrive at

$$\sup_{0<x_n<\rho/2}\frac{v(0,x_n)}{x_n} \leqslant N_5(n,\nu,\sigma,\mathfrak{B})\big(\rho^{-1}\sup_{\mathcal{P}_{\rho/2}\cap\Omega} v + \mathfrak{F}\mathcal{J}_\sigma(\rho/\rho_*)\big). \tag{24}$$

Further, it is easy to find a majorant for $v(0,x_n)/x_n$ for any $x_n \in [\rho/2,\rho)$ since

$$\sup_{\rho/2\leqslant x_n<\rho}\frac{v(0,x_n)}{x_n} \leqslant 2\rho^{-1}\sup_{\rho/2\leqslant x_n<\rho} v(0,x_n) \leqslant 2\rho^{-1}\sup_{\mathcal{P}_\rho\cap\Omega} v. \tag{25}$$

Combining (24) and (25) implies (19) with $C_4 = \max\{N_5,2\}$ for $\rho\leqslant\bar{\rho}$.

Now, we consider $\rho>\bar{\rho}$. If $x_n<\bar{\rho}$ then the estimate

$$\frac{v(0,x_n)}{x_n} \leqslant 2N_5\big(\bar{\rho}^{-1}\sup_{\mathcal{P}_\rho\cap\Omega} v + \mathfrak{F}\mathcal{J}_\sigma(\rho/\rho_*)\big) \tag{26}$$

follows from the above arguments. Otherwise, i.e., for $x_n\geqslant\bar{\rho}$, inequality (26) is especially true. Thus, for $\rho>\bar{\rho}$, we again arrive at (19) with $C_4 = \max\{N_5,2\}\bar{\rho}^{-1}$. □

4. Main results

Recall that Ω satisfies the assumptions on page 442. Throughout this section, we shall suppose that $\mathcal{L}$ is defined by (1), $|\boldsymbol{b}| \in \mathcal{X}(\Omega)$, and a function u satisfies the assumptions

$$u \in \mathcal{W}^2_{\mathcal{X},\mathrm{loc}}(\Omega) \cap \mathcal{C}(\overline{\Omega}), \qquad \mathcal{L}u = 0 \quad \text{in } \Omega, \qquad u|_{\partial\Omega \cap \overline{\mathcal{P}}_{\mathcal{R}_0}} = 0. \tag{27}$$

Theorem 4.1. *Let the inequality*

$$\sup_{x \in \mathcal{P}_{\mathcal{R}_0/2}} \|b^n\|_{\mathcal{X},\mathcal{P}_\rho(x')\cap\Omega} \leqslant \mathfrak{B}\sigma(\rho/\mathcal{R}_0)$$

hold true for all $\rho \leqslant \frac{1}{2}\mathcal{R}_0$. Here $\mathfrak{B}$ is a positive constant, and a function $\sigma \in \mathcal{D}_1$ satisfies

$$\mathcal{J}_\sigma(t) = o(\delta(t)) \quad \text{as } t \to 0. \tag{28}$$

Then, there exists a sufficiently small positive number R_0 completely defined by n, ν, $\mathcal{R}_0$, $\mathfrak{B}$, by the functions σ, δ, and by the moduli of continuity of $|\boldsymbol{b}'|$ in $\mathcal{X}(\Omega)$ such that for any $r \in \left(0, \frac{1}{2}R_0\right)$, we have

$$\operatorname*{osc}_{\Omega\cap\mathcal{P}_{r/4}} \frac{u(x)}{x_n} \leqslant (1 - \varkappa\delta(r)) \operatorname*{osc}_{\Omega\cap\mathcal{P}_{2r}} \frac{u(x)}{x_n}. \tag{29}$$

Here the constant $\varkappa \in (0,1)$ is completely determined by n, ν.

Proof. The proof will be divided into 3 steps.

Step 1: Our arguments are adapted from [Apushkinskaya and Ural′tseva 1995, Lemma 2; Ural′tseva 1996, Lemma 3]. Let us denote

$$m^\pm = \sup_{\Omega\cap\mathcal{P}_{2r}} \pm\frac{u(x)}{x_n}, \qquad \omega = m^+ + m^- = \operatorname*{osc}_{\Omega\cap\mathcal{P}_{2r}} \frac{u(x)}{x_n}.$$

Since $u|_{\partial\Omega} = 0$, we have $m^\pm \geqslant 0$. Therefore, at least one of the numbers $m^\pm$ is not less than $\frac{1}{2}\omega$, and both of the numbers $m^\pm$ are less than ω.

Let $m^+ \geqslant \frac{1}{2}\omega$ for definiteness. Then we consider the nonnegative function $v(x) = m^+x_n - u(x)$ in $\Omega \cap \mathcal{P}_{2r}$; if $m^- > \frac{1}{2}\omega$ then we consider the function $v(x) = m^-x_n + u(x)$.

Due to the definition of δ, for any sufficiently small $r > 0$, we can find a point $x^* \in \partial\mathcal{P}_r \cap \partial\Omega$ such that $x^*_n = r\delta(r)$. Without loss of generality, we may assume that $x^*_1 = r$ and $x^*_\tau = 0$ for $\tau = 2, \dots, n-1$.

Next we assign to x^* a local orthogonal coordinate system $y_1, \dots, y_n$ such that

(a) the y_1-axis is directed along the projection of the vector $(x^*_1, \dots, x^*_{n-1})$ onto a tangential hyperplane to $\partial\Omega$ at x^*;

(b) the y_2-, …, y_{n-1}-axes are parallel to the x_2-, …, x_{n-1}-axes, respectively;

(c) the y_n-axis is directed inside Ω.

Due to the extremal property of x^*, the axes $y_1, \dots, y_{n-1}$ lie in the supporting hyperplane to $\partial\Omega$ at x^*. Moreover, if x^* is a smooth point of $\partial\Omega$ then y_n is directed along the inward normal to $\partial\Omega$.

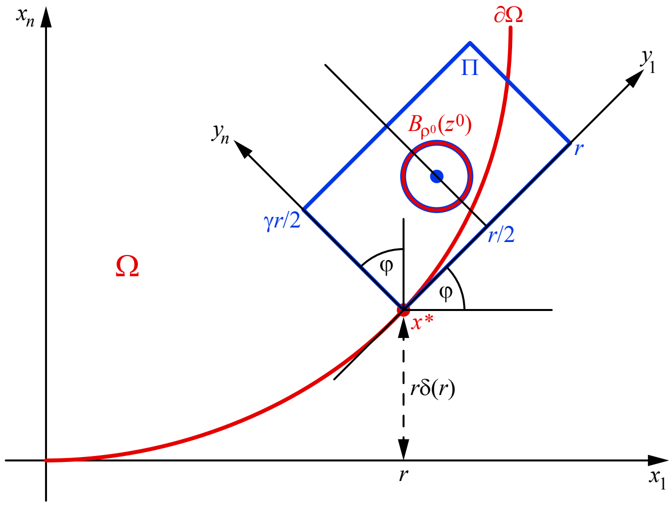


Figure 1. Schematic view of Π and $B_{\rho_0}(z^0)$.

Setting $\gamma = \nu/\sqrt{n-1}$, we consider in y-coordinates the cylinder

$$\Pi := \left\{ y \in \mathbb{R}^n : \left| y_1 - \tfrac{1}{2} r \right| < \tfrac{1}{2} r,\ |y_\tau| < \tfrac{1}{2} r,\ 0 < y_n < \tfrac{1}{2}\gamma r \right\},$$

and the ball $B_{\rho_0}(z^0)$ with $\rho_0 = \frac{1}{8}\gamma r$ and $z^0 = \left(\frac{1}{2}r, 0, \dots, 0, \frac{1}{4}\gamma r\right)$.

It should be emphasized that from now on, all considerations will be carried out in x-coordinates.

We claim that

$$B_{\rho_0}(z^0) \subset \Omega. \tag{30}$$

Indeed, assume that (30) fails. Then there is a point $\hat{x} \in B_{\rho_0}(z^0)$ satisfying (in x-coordinates) the inequalities

$$F(\hat{x}') \geqslant \hat{x}_n \geqslant z_n^0 - \rho_0. \tag{31}$$

Since $\hat{x} \in B_{\rho_0}(z^0)$, it is clear that $|\hat{x}'| \leqslant 2r$ and

$$F(\hat{x}') \leqslant 2r\delta(2r).$$

On the other hand, denoting by φ the angle between the x_n- and y_n-axes (see Figure 1), we conclude that

$$z_n^0 - \rho_0 = r\delta(r) + \tfrac{1}{2} r \sin\varphi + \tfrac{1}{4}\gamma r \cos\varphi - \tfrac{1}{8}\gamma r \geqslant \tfrac{1}{8}\gamma r (2\cos\varphi - 1).$$

Thus (31) is transformed into

$$\gamma(2\cos\varphi - 1) \leqslant 16\delta(2r). \tag{32}$$

In view of (10) and Lemma 2.1, one can choose R_0 so small that $\delta_1(R_0) \leqslant \frac{3}{4}$. It guarantees for all $r \leqslant \frac{1}{2}R_0$, the inequalities

$$\cos\varphi = \frac{1}{\sqrt{1+\tan^2\varphi}} \geqslant \frac{1}{\sqrt{1+\delta_1^2(r)}} \geqslant \frac{1}{\sqrt{1+\delta_1^2(R_0)}} \geqslant \frac{4}{5}. \tag{33}$$

Now, combining (33) and (32), we get a contradiction with relation (10) provided $\delta(R_0)$ is small enough. The proof of (30) is complete.

Step 2: With (30) at hand, we observe that

$$\inf\{x_n \,:\, x \in \Omega \cap \Pi\} \geqslant r\delta(r).$$

On the other hand, the condition $u = 0$ for $x \in \partial\Omega \cap \Pi$ gives the estimate

$$v = m^+ x_n \geqslant \tfrac{1}{2}\omega x_n \quad \text{on } \ \partial\Omega \cap \Pi.$$

Hence,

$$v \geqslant \tfrac{1}{2}\omega r\delta(r) =: k_0 \quad \text{on } \ \partial\Omega \cap \Pi. \tag{34}$$

So, we can apply Lemma 3.1 to the function v in cylinder Π. This gives the estimate

$$\inf_{B_{\rho_0}(z^0)} v \geqslant \big(k_0\big(C_1 - C_2\|\boldsymbol{b}\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big) - C_3\omega r\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+,$$

where C_1, C_2 and C_3 are the constants from Lemma 3.1. Decreasing R_0, if necessary, we may assume that $\|\boldsymbol{b}\|_{\mathcal{X},\Omega\cap\mathcal{P}_{R_0}} \leqslant C_1/(2C_2)$. Thus, we arrive at

$$\inf_{B_{\rho_0}(z^0)} v \geqslant \big(k_0\tfrac{1}{2}C_1 - C_3\omega r\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+ =: k_1. \tag{35}$$

Consider now an arbitrary point $\tilde{z} = (\tilde{z}', \frac{1}{4}r + \frac{1}{8}\rho_0)$ such that $|\tilde{z}'| \leqslant \frac{1}{4}r$. Observe also that $B_{\rho_0}(\tilde{z}) \subset \Omega$, otherwise we get a contradiction with the definition of $\delta(r)$.

We claim that

$$\inf_{B_{\rho_0/8}(\tilde{z})} v \geqslant \big(k_0\widetilde{C}_1 - \widetilde{C}_2\omega r\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+, \tag{36}$$

where $\widetilde{C}_1 = \widetilde{C}_1(n,\nu)$, whereas $\widetilde{C}_2$ is determined completely by n, ν, and $\|\boldsymbol{b}\|_{\mathcal{X},\Omega}$. Indeed, due to the convexity of Ω, for l running from 1 to a finite number $\mathfrak{N} = \mathfrak{N}(n,\nu)$ chosen so that

$$\frac{4}{3\rho_0}|z^0 - \tilde{z}| \leqslant \mathfrak{N} \leqslant \frac{2}{\rho_0}|z^0 - \tilde{z}|, \tag{37}$$

and for points $z^{[l]} := z^0 - (l/\mathfrak{N})(z^0 - \tilde{z})$, we have $B_{\rho_0}(z^{[l]}) \subset \Omega$. It should be emphasized that the lower and the upper bounds in (37) do not depend on r.

In view of (35), we can compare in $\mathcal{B}(z^{[1]}, \frac{1}{8}\rho_0, \rho_0)$ the function v with the standard barrier function

$$w(x) = k_1\,\frac{|x - z^{[1]}|^{-s} - \rho_0^{-s}}{\big(\frac{1}{8}\rho_0\big)^{-s} - \rho_0^{-s}}.$$

If $s = n\nu^{-2}$ then elementary calculation guarantees the estimates

$$\begin{aligned} \mathcal{L}w \leqslant |\boldsymbol{b}||Dw| \leqslant c(n,\nu)k_1|\boldsymbol{b}|\rho_0^{-1} \quad & \text{in } \ \mathcal{B}(z^{[1]}, \tfrac{1}{8}\rho_0, \rho_0), \\ w(x) = k_1 \leqslant v(x) \quad & \text{on the sphere } \ |x - z^{[1]}| = \tfrac{1}{8}\rho_0, \\ w(x) = 0 \leqslant v(x) \quad & \text{on the sphere } \ |x - z^{[1]}| = \rho_0. \end{aligned}$$

Applying the maximum principle (11) in $\mathcal{B}\big(z^{[1]}, \frac{1}{8}\rho_0, \rho_0\big)$ to the difference $w - v$ gives us the inequality

$$v(x) \geqslant \big(k_1\big(w(x) - 2cN_0\|\boldsymbol{b}\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big) - N_0\tfrac{1}{4}\gamma r\omega\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+.$$

Since $B_{\rho_0/8}(z^{[2]}) \subset \mathcal{B}\big(z^{[1]}, \frac{1}{8}\rho_0, \frac{7}{8}\rho_0\big)$, the evident bound $w \geqslant \theta(n,\nu)$ holds true in $B_{\rho_0/8}(z^{[2]})$.

Decreasing R_0, if necessary, we ensure that $\|\boldsymbol{b}\|_{\mathcal{X},\Omega\cap\mathcal{P}_{R_0}} \leqslant (4cN_0)^{-1}\theta$. This implies

$$\inf_{B_{\rho_0/8}(z^{[2]})} v(x) \geqslant \big(\tfrac{1}{2}k_1\theta - N_0\tfrac{1}{4}\gamma r\omega\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+ =: k_2.$$

Repeating this procedure for $\mathcal{B}\big(z^{[l]}, \frac{1}{8}\rho_0, \rho_0\big)$ and $l = 2, \dots, \mathfrak{N}$, we arrive at (36) with $\widetilde{C}_1 = \big(\frac{1}{2}\theta\big)^{\mathfrak{N}}$ and

$$\widetilde{C}_2 = N_0\tfrac{1}{4}\gamma \cdot \frac{1-\big(\frac{1}{2}\theta\big)^{\mathfrak{N}}}{1-\frac{1}{2}\theta}.$$

Furthermore, it is clear that

$$\big(k_0\widetilde{C}_1 - \widetilde{C}_2 r\omega\|b^n\|_{\mathcal{X},\Omega\cap\mathcal{P}_{2r}}\big)_+ \geqslant \omega r\big(\tfrac{1}{2}\widetilde{C}_1\delta(r) - \widetilde{C}_2\mathfrak{B}\sigma(r/\mathcal{R}_0)\big)_+,$$

while inequalities (3) and (4) guarantee that

$$\sigma(r/\mathcal{R}_0) \leqslant \frac{\mathcal{J}_\sigma(r)}{\mathcal{R}_0}.$$

Decreasing R_0 again and taking into account the assumption (28) and the above inequalities, we can transform (36) into the form

$$\inf_{B_{\rho_0/8}(\tilde{z})} v \geqslant \tfrac{1}{4}\widetilde{C}_1\omega r\delta(r) =: \tilde{k}. \tag{38}$$

Step 3: Now, we take a small $\eta > 0$, define the set

$$\mathcal{A}_\eta := \mathcal{B}\big(\tilde{z}, \tfrac{1}{8}\rho_0, \tilde{z}_n\big) \cap \Omega \cap \big\{x \in \mathcal{P}_{\mathcal{R}_0} \,:\, F(x') + \eta < x_n < \mathcal{R}_0\big\}$$

and introduce in $\mathcal{A}_\eta$ the barrier function

$$W(x) = \mu\tilde{k}\,\frac{|x-\tilde{z}|^{-s} - (\tilde{z}_n)^{-s}}{\big(\frac{1}{8}\rho_0\big)^{-s} - (\tilde{z}_n)^{-s}},$$

where $s = n\nu^{-2}$ and $0 < \mu \leqslant 1$.

Notice that $D(Du) \in \mathcal{X}(\mathcal{A}_\eta)$. Using Lemma 2.8, we construct the family of operators $\mathcal{L}_\varepsilon$ satisfying $\|\mathcal{L}_\varepsilon u\|_{\mathcal{X},\mathcal{A}_\eta} \to 0$ as $\varepsilon \to 0$.

Arguing in the spirit of the proof of Lemma 4.2 [Ladyzhenskaya and Ural′tseva 1988], we define $v_1(x)$ and $v_2(x)$ as solutions of the problems

$$\begin{cases} \mathcal{L}_\varepsilon v_1 = b^i_\varepsilon D_i W & \text{in } \mathcal{A}_\eta, \\ \quad v_1 = v & \text{on } \partial\mathcal{A}_\eta, \end{cases} \qquad \begin{cases} \mathcal{L}_\varepsilon v_2 = b^i_\varepsilon D_i W - b^n_\varepsilon m^+ & \text{in } \mathcal{A}_\eta, \\ \quad v_2 = 0 & \text{on } \partial\mathcal{A}_\eta. \end{cases}$$

It is well known (see, for instance, [Krylov 2008, Chapter 6]) that $D(Dv_1)$ and $D(Dv_2)$ belong to the space $\mathrm{BMO}_{\mathrm{loc}}(\mathcal{A}_\eta)$. Moreover, the John–Nirenberg theorem [1961] (see also [Duoandikoetxea 2001, §4, Chapter 6]) implies that $D(Dv_i)$, with $i = 1, 2$, belong to the Orlicz space $L_{\Phi,\mathrm{loc}}(\mathcal{A}_\eta)$ with $\Phi(\xi) = e^\xi - \xi - 1$. So, taking into account the property (iii), we may conclude that $v_i \in \mathcal{W}^2_{\mathcal{X},\mathrm{loc}}(\mathcal{A}_\eta)$, with $i = 1, 2$.

Furthermore, in view of (38) and by direct calculation, we have the inequalities

$$\begin{aligned} \mathcal{L}_\varepsilon W &\leqslant b^i_\varepsilon D_i W && \text{in } \mathcal{A}_\eta, \\ W(x) = \mu \tilde{k} &\leqslant v(x) = v_1(x) && \text{on the sphere } |x-\tilde{z}| = \tfrac{1}{8}\rho_0, \\ W(x) = 0 &\leqslant v(x) = v_1(x) && \text{on } \partial\mathcal{A}_\eta \cap \{x \in \mathbb{R}^n : |x-\tilde{z}| = \tilde{z}_n\}. \end{aligned}$$

On the rest of $\partial\mathcal{A}_\eta$, we have $x_n = F(x') + \eta$ and, consequently, $\operatorname{dist}\{x, \partial\Omega\} \leqslant \eta$. Since $u \in \mathcal{C}(\overline{\Omega})$, the latter inequality implies the estimate $u \leqslant H(\eta)$ there, and therefore,

$$v_1(x) = v(x) = m^+ x_n - u \geqslant \tfrac{1}{2}\omega x_n - H(\eta),$$

where H is a nonnegative function tending to zero as $\eta \to 0$.

In addition, it is easy to verify that

$$W(x) \leqslant \mu N_6(n,\nu)\widetilde{C}_1 \omega\, \delta(r) x_n \quad \text{in } \overline{\mathcal{B}}\big(\tilde{z}, \tfrac{1}{8}\rho_0, \tilde{z}_n\big).$$

Choosing $\mu = \min\{1, (2N_6\widetilde{C}_1)^{-1}\}$, we get

$$v_1(x) \geqslant W(x) - H(\eta) \quad \text{on } \partial\mathcal{A}_\eta.$$

The maximum principle (11) applied to the difference $W - H(\eta) - v_1$ in $\mathcal{A}_\eta$ provides the inequality

$$v_1(x) \geqslant W(x) - H(\eta) \geqslant \mu N_7(n,\nu)\widetilde{C}_1 \omega\, \delta(r)(\tilde{z}_n - |x-\tilde{z}|) - H(\eta).$$

It follows from the last inequality with $x = (\tilde{z}', x_n) \in \Omega$ and $0 < x_n \leqslant \tilde{z}_n - \frac{1}{8}\rho_0 = \frac{1}{4}r$ that

$$v_1(\tilde{z}', x_n) \geqslant N_8(n,\nu)\, \omega\, \delta(r) x_n - H(\eta). \tag{39}$$

Next, we look for a majorant for v_2. With this aim in view, we extend the coefficients a^{ij}_ε continuously and the coefficients b^i_ε by zero to the whole annulus $\mathcal{B}\big(\tilde{z}, \frac{1}{8}\rho_0, \tilde{z}_n\big)$, and denote by $\tilde{v}_2(x)$ the solution of the problem

$$\mathcal{L}_\varepsilon \tilde{v}_2 = \begin{cases} (\mathcal{L}_\varepsilon v_2)_+ & \text{in } \mathcal{A}_\eta, \\ 0 & \text{in } \mathcal{B}\big(\tilde{z}, \tfrac{1}{8}\rho_0, \tilde{z}_n\big) \setminus \mathcal{A}_\eta, \end{cases}$$

$$\tilde{v}_2 = 0 \quad \text{on } \partial\mathcal{B}\big(\tilde{z}, \tfrac{1}{8}\rho_0, \tilde{z}_n\big).$$

The maximum principle guarantees

$$v_2 \leqslant \tilde{v}_2 \quad \text{in } \mathcal{A}_\eta. \tag{40}$$

Direct computations show that for $\rho \leqslant \frac{1}{4}r$ the barrier function W satisfies in the set

$$\mathcal{E}_\rho := \mathcal{P}_\rho(\tilde{z}', 0) \cap \mathcal{B}\big(\tilde{z}, \tfrac{1}{8}\rho_0, \tilde{z}_n\big)$$

the following inequalities

$$|D_n W| \leqslant |DW| \leqslant N_9(n,\nu)\, \mu \frac{\tilde{k}}{r} \leqslant N_9\, \omega\, \delta(r),$$

$$|D'W| \leqslant N_9 \mu \frac{\tilde{k}\rho}{r^2} \leqslant N_9\, \omega \frac{\delta(r)\rho}{r}.$$

So, in view of (15) and (10), we have for all $\rho \leqslant \frac{1}{4}r$, the bounds

$$\begin{aligned}\|(\mathcal{L}_\varepsilon \tilde{v}_2)_+\|_{\mathcal{X},\mathcal{E}_\rho} &\leqslant \|b^n\|_{\mathcal{X},\mathcal{E}_\rho}\big(m^+ + \|D_n W\|_{\infty,\mathcal{E}_\rho}\big) + \|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{E}_\rho}\|D'W\|_{\infty,\mathcal{E}_\rho}\\ &\leqslant N_{10}(n,\nu)\,\omega\left(\mathfrak{B}\sigma\left(\frac{\rho}{\mathcal{R}_0}\right) + \frac{\delta(r)}{r}\rho\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta}\right).\end{aligned}$$

Since the function $\rho \mapsto \big(\mathfrak{B}\sigma(\rho/\mathcal{R}_0) + (\delta(r)/r)\rho\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta}\big)$ satisfies the Dini condition at zero, there exist the uniquely defined function $\sigma_1 \in \mathcal{D}_1$ and a constant $\mathfrak{B}_1$ such that

$$\mathfrak{B}\sigma\left(\frac{\rho}{\mathcal{R}_0}\right) + \frac{\delta(r)}{r}\rho\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta} = \mathfrak{B}_1\sigma_1\left(\frac{4\rho}{r}\right).$$

Thus, we may apply Lemma 3.2 to the function $\tilde{v}_2$. It gives for $\rho = \frac{1}{4}r$ the estimate

$$\sup_{0<x_n<r/4} \frac{\tilde{v}_2(\tilde{z}', x_n)}{x_n} \leqslant C_4\Big(\big(\tfrac{1}{4}r\big)^{-1} \sup_{\mathcal{E}_{r/4}} \tilde{v}_2 + N_{10}\omega\mathfrak{B}_1\mathcal{J}_{\sigma_1}(1)\Big). \tag{41}$$

It is easy to see that

$$\mathfrak{B}_1\mathcal{J}_{\sigma_1}(1) = \mathfrak{B}\mathcal{J}_\sigma\left(\frac{r}{4\mathcal{R}_0}\right) + \tfrac{1}{4}\delta(r)\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta}.$$

Furthermore, applying (11) to $\tilde{v}_2$ and to the operator $\mathcal{L}_\varepsilon$ in $\mathcal{B}\big(\tilde{z}, \frac{1}{8}\rho_0, \tilde{z}_n\big)$, we obtain

$$\sup_{\mathcal{E}_{r/4}} \tilde{v}_2 \leqslant \sup_{\mathcal{B}(\tilde{z},\rho_0/8,\tilde{z}_n)} \tilde{v}_2 \leqslant N_{11}(n,\nu,\|\boldsymbol{b}\|_{\mathcal{X},\Omega})\,\omega r\left(\mathfrak{B}\sigma\left(\frac{r}{\mathcal{R}_0}\right) + \delta(r)\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta}\right).$$

Substitution of the above estimates in (41) and consideration of (3) provide

$$\sup_{0<x_n<r/4} \frac{\tilde{v}_2(\tilde{z}', x_n)}{x_n} \leqslant N_{12}\,\omega\left(\mathfrak{B}\mathcal{J}_\sigma\left(\frac{r}{\mathcal{R}_0}\right) + \delta(r)\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta}\right), \tag{42}$$

where the constant N_{12} depends only on n, ν and $\|\boldsymbol{b}\|_{\mathcal{X},\Omega}$.

Taking into account the inequality (5), the assumption (28), and the evident relation $\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}} = o(1)$ as $r \to 0$, we decrease R_0 such that the property

$$\mathfrak{B}\mathcal{J}_\sigma\left(\frac{r}{\mathcal{R}_0}\right) + \delta(r)\|\boldsymbol{b}'\|_{\mathcal{X},\mathcal{A}_\eta} \leqslant \frac{N_8}{2N_{12}}\delta(r) \tag{43}$$

holds true for all $r \leqslant R_0$.

Finally, combining (39)–(40) with (42)–(43), we arrive at the estimate

$$v_1(\tilde{z}', x_n) - v_2(\tilde{z}', x_n) \geqslant \tfrac{1}{2}N_8\omega\,\delta(r)x_n - H(\eta) \tag{44}$$

for $r \leqslant R_0$ and $x = (\tilde{z}', x_n) \in \Omega$ with $x_n \in \big[F(\tilde{z}') + \eta, \frac{1}{4}r\big]$.

Considering in $\mathcal{A}_\eta$ the function $v_3(x) = v(x) - v_1(x) + v_2(x)$, one can easily see that

$$\mathcal{L}_\varepsilon v_3 = -\mathcal{L}_\varepsilon u \to 0 \quad \text{in } \mathcal{X}(\mathcal{A}_\eta) \text{ as } \varepsilon \to 0.$$

In addition, $v_3 = 0$ on $\partial\mathcal{A}_\eta$. Applying the maximum principle (11) to $\pm v_3$ and to the operator $\mathcal{L}_\varepsilon$, we obtain that the difference $v_1(x) - v_2(x)$ converges to $v(x)$ uniformly in $\mathcal{A}_\eta$. Therefore, passing in (44) first to the limit as $\varepsilon \to 0$ and then as $\eta \to 0$, we get

$$\frac{v(x)}{x_n} \geqslant \tfrac{1}{2} N_8 \omega \delta(r) \tag{45}$$

for $r \leqslant R_0$ and $x = (\tilde{z}', x_n) \in \Omega$ with $x_n \in [F(\tilde{z}'), \frac{1}{4}r]$.

Since $\tilde{z}'$ can be chosen arbitrarily with only $|\tilde{z}'| \leqslant \frac{1}{4}r$, the estimate (45) gives (29) with $\varkappa = \frac{1}{2}N_8$. □

Theorem 4.2 (main theorem). *Let the assumptions of Theorem 4.1 hold*, *and suppose*

$$\delta(r) = \max_{|x'|\leqslant r} \frac{F(x')}{|x'|}$$

does not satisfy the Dini condition at zero.

Then for any function u satisfying (27), *the equality*

$$\frac{\partial u}{\partial \boldsymbol{n}}(0) = 0$$

holds true.

Proof. Consider the sequence $r_k = 8^{-k} R_0$, with $k \geqslant 0$, where R_0 is the constant from Theorem 4.1.

Applying Theorem 4.1 to u guarantees for $k \geqslant 0$ the inequalities

$$\operatorname*{osc}_{\Omega\cap\mathcal{P}_{r_{k+1}}} \frac{u(x)}{x_n} \leqslant \bigl(1 - \varkappa\delta\bigl(\tfrac{1}{2}r_k\bigr)\bigr) \operatorname*{osc}_{\Omega\cap\mathcal{P}_{r_k}} \frac{u(x)}{x_n} \leqslant \operatorname*{osc}_{\Omega\cap\mathcal{P}_{R_0}} \frac{u(x)}{x_n} \cdot \prod_{j=0}^{k} \bigl(1 - \varkappa\delta\bigl(\tfrac{1}{2}r_j\bigr)\bigr).$$

Since

$$\sum_{j=0}^{\infty} \ln\bigl(1 - \varkappa\delta\bigl(\tfrac{1}{2}r_j\bigr)\bigr) \asymp -\sum_{j=0}^{\infty} \delta\bigl(\tfrac{1}{2}r_j\bigr) \asymp -\int_0^{r_0} \frac{\delta(r)}{r}\, dr = -\infty,$$

we have

$$\prod_{j=0}^{k} \bigl(1 - \varkappa\delta\bigl(\tfrac{1}{2}r_j\bigr)\bigr) \to 0 \quad \text{as } k \to \infty.$$

We recall also that Lemma 3.2 implies the finiteness of the quantity $\operatorname*{osc}_{\Omega\cap\mathcal{P}_{R_0}} (u(x)/x_n)$.

Thus, taking into account that $u|_{\partial\Omega\cap\mathcal{P}_{\mathcal{R}_0}} = 0$, we get

$$\left|\frac{\partial u}{\partial \boldsymbol{n}}(0)\right| = \left|\lim_{x_n\to 0} \frac{u(0, x_n)}{x_n}\right| \leqslant \lim_{k\to\infty} \left|\operatorname*{osc}_{\Omega\cap\mathcal{P}_{r_k}} \frac{u(x)}{x_n}\right| = 0.$$ □

Acknowledgements

The authors would like to thank the anonymous referee for essential comments.

This work was supported by the Russian Foundation of Basic Research (RFBR) through the grant 15-01-07650 and by the St. Petersburg State University grant 6.38.670.2013.

References

[Aleksandrov 1960] A. D. Aleksandrov, "Certain estimates for the Dirichlet problem", *Dokl. Akad. Nauk SSSR* **134** (1960), 1001–1004. In Russian; translated in *Soviet Math. Dokl.* **1** (1961), 1151–1154. MR 26 #5290 Zbl 0100.09303

[Aleksandrov 1963] A. D. Aleksandrov, "Uniqueness conditions and bounds for the solution of the Dirichlet problem", *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom.* **18**:3 (1963), 5–29. In Russian. MR 29 #1434

[Alvarado 2011] R. Alvarado, *Topics in harmonic analysis and partial differential equations: extension theorems and geometric maximum principles*, master's thesis, University of Missouri, 2011, available at http://tinyurl.com/hzgwt8t.

[Alvarado et al. 2011] R. Alvarado, D. Brigham, V. Maz'ya, M. Mitrea, and E. Ziadé, "On the regularity of domains satisfying a uniform hour-glass condition and a sharp version of the Hopf–Oleinik boundary point principle", *Problem. Mat. Anal.* **57** (2011), 3–68. In Russian; translated in *J. Math. Sci. New York* **176**:3 (2011), 281–360. MR 2012j:35060 Zbl 1290.35046

[Apushkinskaya and Nazarov 2000] D. E. Apushkinskaya and A. I. Nazarov, "The Dirichlet problem for quasilinear elliptic equations in domains with smooth closed edges", *Problem. Mat. Anal.* **21** (2000), 3–29. In Russian; translated in *J. Math. Sci. New York* **105**:5 (2001), 2299–2318. MR 2002m:35061 Zbl 0987.35043

[Apushkinskaya and Ural′tseva 1995] D. E. Apushkinskaya and N. N. Ural′tseva, "On the behavior of the free boundary near the boundary of the domain", *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **221** (1995), 5–19, 253. In Russian; translated in *J. Math. Sci. (New York)* **87**:2 (1997), 3267–3276. MR 96m:35340 Zbl 0888.35141

[Bakel′man 1961] I. J. Bakel′man, "On the theory of quasilinear elliptic equations", *Sibirsk. Mat. Ž.* **2** (1961), 179–186. In Russian. MR 23 #A3900 Zbl 0100.30503

[Duoandikoetxea 2001] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics **29**, American Mathematical Society, Providence, RI, 2001. MR 2001k:42001 Zbl 0969.42001

[Giraud 1932] G. Giraud, "Generalisation des problèmes sur les operations du type elliptique", *Bull. Sci, Math.* (2) **56** (1932), 316–352. JFM 58.0494.02

[Giraud 1933] G. Giraud, "Problèmes de valeurs à la frontière relatifs à certaines données discontinues", *Bull. Soc. Math. France* **61** (1933), 1–54. MR 1504997 Zbl 0007.11502

[Himčenko 1970] B. N. Himčenko, "On the behavior of solutions of elliptic equations near the boundary of a domain of type $A^{(1)}$", *Dokl. Akad. Nauk SSSR* **193** (1970), 304–305. In Russian; translated in *Soviet Math. Dokl.* **11** (1970), 943–944. MR 42 #8071 Zbl 0206.40604

[Hopf 1952] E. Hopf, "A remark on linear elliptic differential equations of second order", *Proc. Amer. Math. Soc.* **3** (1952), 791–793. MR 14,280b Zbl 0048.07802

[John and Nirenberg 1961] F. John and L. Nirenberg, "On functions of bounded mean oscillation", *Comm. Pure Appl. Math.* **14** (1961), 415–426. MR 24 #A1348 Zbl 0102.04302

[Kamynin and Himčenko 1975] L. I. Kamynin and B. N. Himčenko, "Theorems of Giraud type for a second order elliptic operator that is weakly degenerate near the boundary", *Dokl. Akad. Nauk SSSR* **224**:4 (1975), 752–755. In Russian; translated in *Soviet Math. Dokl.* **16**:5 (1975), 1287–1291. MR 53 #1019 Zbl 0345.35046

[Kamynin and Himčenko 1977] L. I. Kamynin and B. N. Himčenko, "Theorems of Giraud type for second order equations with a weakly degenerate non-negative characteristic part", *Sibirsk. Mat. Ž.* **18**:1 (1977), 103–121, 238. In Russian; translated in *Sib. Math. J.* **18** (1977), 76–91. MR 55 #10831 Zbl 374.35013

[Kantorovich and Akilov 1982] L. V. Kantorovich and G. P. Akilov, *Functional analysis*, 2nd ed., Pergamon Press, Oxford-Elmsford, N.Y., 1982. MR 83h:46002 Zbl 0484.46003

[Keldysh and Lavrent'ev 1937] M. V. Keldysh and M. A. Lavrent'ev, "On the uniqueness of the Neumann problem", *Dokl. Akad. Nauk SSSR* **16**:3 (1937), 151–152. In Russian.

[Korn 1901] A. Korn, *Lehrbuch der Potentialtheorie, II: Allgemeine Theorie des Logarithmischen Potentials und der Potentialfunctionen in der Ebene*, F. Dümmler, Berlin, 1901. JFM 31.0728.04

[Krylov 2008] N. V. Krylov, *Lectures on elliptic and parabolic equations in Sobolev spaces*, Graduate Studies in Mathematics **96**, American Mathematical Society, Providence, RI, 2008. MR 2009k:35001 Zbl 1147.35001

[Ladyzhenskaya and Ural′tseva 1985] O. A. Ladyzhenskaya and N. N. Ural′tseva, "Estimates of the Hölder constant for functions satisfying a uniformly elliptic or uniformly parabolic quasilinear inequality with unbounded coefficients", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **147** (1985), 72–94, 204. In Russian; translated in *J. Soviet Math.* **37**:1 (1987), 837–851. MR 87e:35090 Zbl 0596.35128

[Ladyzhenskaya and Ural′tseva 1988] O. A. Ladyzhenskaya and N. N. Ural′tseva, "Estimates on the boundary of the domain of first derivatives of functions satisfying an elliptic or a parabolic inequality", *Trudy Mat. Inst. Steklov.* **179** (1988), 102–125, 243. In Russian; translated in *Proc. Steklov Inst. Math.* **179**, (1989), 109–135. MR 89k:35083 Zbl 0707.35023

[Lichtenstein 1924] L. Lichtenstein, "Neue Beiträge zur Theorie der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus", *Math. Z.* **20**:1 (1924), 194–212. MR 1544673

[Lieberman 1985] G. M. Lieberman, "Regularized distance and its applications", *Pacific J. Math.* **117**:2 (1985), 329–352. MR 87j:35101 Zbl 0535.35028

[Mikayelyan and Shahgholian 2015] H. Mikayelyan and H. Shahgholian, "Hopf's lemma for a class of singular/degenerate PDEs", *Ann. Acad. Sci. Fenn. Math.* **40**:1 (2015), 475–484. MR 3329156 Zbl 1326.35156

[Nadirashvili 1983] N. S. Nadirashvili, "On the question of the uniqueness of the solution of the second boundary value problem for second-order elliptic equations", *Mat. Sb.* (*N.S.*) **122(164)**:3 (1983), 341–359. In Russian; translated in *Math. USSR-Sb.* **50**:2 (1985), 325–341. MR 85f:35070 Zbl 0563.35026

[Nazarov 2001] A. I. Nazarov, "Estimates for the maximum of solutions of elliptic and parabolic equations in terms of weighted norms of the right-hand side", *Algebra i Analiz* **13**:2 (2001), 151–164. In Russian; translated in *St. Petersburg Math. J.* **13**:2 (2002), 269–279. MR 2002j:35049 Zbl 0994.35032

[Nazarov 2005] A. I. Nazarov, "The maximum principle of A. D. Aleksandrov", *Sovrem. Mat. Prilozh.* 29 (2005), 129–145. In Russian; translated in *J. Math. Sci. New York* **142**:3 (2007), 2154–2171. MR 2011b:35051 Zbl 1157.35019

[Nazarov 2012] A. I. Nazarov, "A centennial of the Zaremba–Hopf–Oleinik lemma", *SIAM J. Math. Anal.* **44**:1 (2012), 437–453. MR 2888295 Zbl 1243.35031

[Nazarov and Uraltseva 2009] A. I. Nazarov and N. N. Uraltseva, "Qualitative properties of solutions to elliptic and parabolic equations with unbounded lower-order coefficients", preprint, 2009, available at http://www.mathsoc.spb.ru/preprint/2009/09-05.ps.gz.

[Neumann 1888] C. Neumann, "Ueber die Methode des arithmetischen Mittels, zweite Abhandlung", *Leipz. Abh.* **14** (1888), 565–726. JFM 20.1015.01

[Oleĭnik 1952] O. A. Oleĭnik, "On properties of solutions of certain boundary problems for equations of elliptic type", *Mat. Sbornik N.S.* **30(72)**:3 (1952), 695–702. In Russian. MR 14,280a Zbl 0046.10401

[Safonov 2008] M. V. Safonov, "Boundary estimates for positive solutions to second order elliptic equations", preprint, 2008. arXiv 0810.0522

[Safonov 2010] M. V. Safonov, "Non-divergence elliptic equations of second order with unbounded drift", pp. 211–232 in *Nonlinear partial differential equations and related topics*, edited by A. A. Arkhipova and A. I. Nazarov, Amer. Math. Soc. Transl. Ser. 2 **229**, Amer. Math. Soc., Providence, RI, 2010. MR 2011i:35040 Zbl 1208.35028

[Sweers 1997] G. Sweers, "Hopf's lemma and two-dimensional domains with corners", *Rend. Istit. Mat. Univ. Trieste* **28**:suppl. (1997), 383–419. MR 99e:35042 Zbl 0892.35032

[Tolksdorf 1983] P. Tolksdorf, "On the Dirichlet problem for quasilinear equations in domains with conical boundary points", *Comm. Partial Differential Equations* **8**:7 (1983), 773–817. MR 85g:35053 Zbl 0515.35024

[Ural′tseva 1996] N. N. Ural′tseva, "C^1 regularity of the boundary of a noncoincident set in a problem with an obstacle", *Algebra i Analiz* **8**:2 (1996), 205–221. In Russian; translated in *St. Petersburg Math. J.* **8**:2 (1997), 241–253. MR 97m:35105 Zbl 0869.35116

[Widman 1967] K.-O. Widman, "Inequalities for the Green function and boundary continuity of the gradient of solutions of elliptic differential equations", *Math. Scand.* **21** (1967), 17–37. MR 39 #621 Zbl 0164.13101

[Zaremba 1910] S. Zaremba, "Sur un problème mixte relatif à l' équation de Laplace", *Krakau Anz. (A)* (1910), 313–344. JFM 41.0854.12

Received 2 Apr 2015. Revised 29 Sep 2015. Accepted 16 Dec 2015.

DARYA E. APUSHKINSKAYA: darya@math.uni-sb.de
Department of Mathematics, Saarland University, P.O. Box 151150, Saarbrücken 66041, Germany
and
St. Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg 199034, Russia

ALEXANDER I. NAZAROV: al.il.nazarov@gmail.com
St. Petersburg Department, Steklov Institute, Fontanka 27, St. Petersburg 191023, Russia
and
St. Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg 199034, Russia

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.459

GROUND STATES OF LARGE BOSONIC SYSTEMS: THE GROSS–PITAEVSKII LIMIT REVISITED

PHAN THÀNH NAM, NICOLAS ROUGERIE AND ROBERT SEIRINGER

We study the ground state of a dilute Bose gas in a scaling limit where the Gross–Pitaevskii functional emerges. This is a repulsive nonlinear Schrödinger functional whose quartic term is proportional to the scattering length of the interparticle interaction potential. We propose a new derivation of this limit problem, with a method that bypasses some of the technical difficulties that previous derivations had to face. The new method is based on a combination of Dyson's lemma, the quantum de Finetti theorem and a second moment estimate for ground states of the effective Dyson Hamiltonian. It applies equally well to the case where magnetic fields or rotation are present.

1. Introduction

The rigorous derivation of effective nonlinear theories from many-body quantum mechanics has been studied extensively in recent years, motivated in part by experiments in cold atom physics. For bosons, the emergence of the limit theories can be interpreted as due to most of the particles occupying the same quantum state: this is the Bose–Einstein condensation phenomenon, observed first in dilute alkali vapors some twenty years ago.

The parameter regime most relevant for the description of the actual physical setup is the Gross–Pitaevskii limit. It is also the most mathematically demanding regime considered in the literature so far; see [Lieb and Yngvason 1998; Lieb et al. 2000; Lieb and Seiringer 2002; 2006] for the derivation of equilibrium states and [Erdős et al. 2009; 2010; Benedikter et al. 2015; Pickl 2015] for dynamics (more extensive lists of references may be found in [Lieb et al. 2005b; Rougerie 2015; Benedikter et al. 2016]). The main reason for this sophistication is the fact that interparticle correlations due to two-body scattering play a leading-order role in this regime. The goal of this paper is to present a method for the derivation of Gross–Pitaevskii theory at the level of the ground state that is conceptually and technically simpler

MSC2010: 35Q40, 81V70.
Keywords: many-body quantum mechanics, mean-field limits, Bose–Einstein condensates.

than existing proofs, in particular that of [Lieb and Seiringer 2006], which was so far the only method applicable when an external magnetic field is present.

Our setting is as follows: we consider N interacting bosons in the three-dimensional space $\mathbb{R}^3$, described by the many-body Schrödinger Hamiltonian

$$H_N = \sum_{j=1}^{N} h_j + \sum_{1 \leqslant j<k \leqslant N} w_N(x_j - x_k) \tag{1-1}$$

acting on the space $\mathfrak{H}^N = \bigotimes_{\text{sym}}^N L^2(\mathbb{R}^3)$ of permutation-symmetric square integrable functions. The one-body operator is given by

$$h := (-i\nabla + A(x))^2 + V(x),$$

with a magnetic (or a rotation) field A satisfying

$$A \in L^3_{\text{loc}}(\mathbb{R}^3, \mathbb{R}^3), \qquad \lim_{|x|\to\infty} |A(x)|e^{-b|x|} = 0, \tag{1-2}$$

for some constant $b > 0$ and an external potential V satisfying

$$0 \leqslant V \in L^1_{\text{loc}}(\mathbb{R}^3), \qquad \lim_{|x|\to\infty} V(x) = +\infty. \tag{1-3}$$

We thus consider nonrelativistic particles in a trapping potential, possibly under the influence of an effective magnetic field, which might be due to rotation of the sample or the interaction with optical fields.

The particles interact pairwise via a repulsive potential w_N given by

$$w_N(x) = N^2 w(Nx), \tag{1-4}$$

where w is a fixed function which is nonnegative, radial and of finite range, i.e., $\mathbb{1}(|x| > R_0)w(x) \equiv 0$ for some constant $R_0 > 0$. Different scalings of the interaction potential of the form

$$w_{\beta,N}(x) = \frac{1}{N} N^{3\beta} w(N^\beta x), \tag{1-5}$$

with $0 \leqslant \beta \leqslant 1$, have been considered in the literature. The N^{-1} prefactor makes the interaction energy in (1-1) of the same order as the one-particle energy. Indeed, if $\beta > 0$, then

$$N^{3\beta} w(N^\beta x) \xrightarrow[N\to\infty]{} \left(\int w\right)\delta_0 \tag{1-6}$$

weakly and thus the interaction potential $w_{\beta,N}$ should be thought of as leading to a bounded interaction energy per pair of particles. Generally speaking, the larger the parameter β, the faster the potential converges to a point interaction, and thus the harder the analysis. Note that the cases $\beta < \frac{1}{3}$ and $\beta > \frac{1}{3}$ correspond to two physically rather different scenarios: in the former, the range of the potential is much larger than the typical interparticle distance $N^{-1/3}$, and we should expect many weak collisions; while in the latter, we rather have very few but very strong collisions. In this paper, we consider the most interesting case $\beta = 1$, where the naive approximation (1-6) does *not* capture the leading-order behavior of the physical system. In fact, the strong correlations at short distances $O(N^{-1})$ yield a nonlinear correction, which essentially amounts to replacing the coupling constant $\int w$ by $(8\pi) \times$ (the scattering length of w).

Let us quickly recall the definition of the scattering length; a more complete discussion can be found in [Lieb et al. 2005b, Appendix C]. Under our assumption on w, the zero-energy scattering equation

$$(-2\Delta + w(x))f(x), \qquad \lim_{|x|\to\infty} f(x) = 1,$$

has a unique solution and it satisfies

$$f(x) = 1 - \frac{a}{|x|} \quad \forall |x| > R_0$$

for some constant $a \geqslant 0$ which is called the *scattering length* of w. In particular, if w is the potential for hard spheres, namely $w(x) \equiv +\infty$ when $|x| < R_0$ and $w(x) \equiv 0$ when $|x| \geqslant R_0$, then the scattering length of w is exactly R_0. In a dilute gas, the scattering length can be interpreted as an effective range of the interaction: a quantum particle far from the others is felt by them as a hard sphere of radius a. A useful variational characterization of a is

$$8\pi a = \inf\left\{ \int_{\mathbb{R}^3} 2|\nabla f|^2 + w|f|^2, \ \lim_{|x|\to\infty} f(x) = 1 \right\}. \tag{1-7}$$

Consequently, $8\pi a$ is smaller than $\int w$ (the strict inequality can be seen by taking the trial function $1-\lambda g$ with $g \in C_c^2(\mathbb{R}^3, \mathbb{R})$ satisfying $g(x) \equiv 1$ when $|x| < R_0$, and $\lambda > 0$ sufficiently small). Moreover, a simple scaling shows that the scattering length of $w_N = N^2 w(N\,\cdot)$ is a/N.

We are going to prove that the ground-state energy and ground states of H_N converge to those of the Gross–Pitaevskii functional

$$\mathcal{E}_{\mathrm{GP}}(u) := \langle u, hu\rangle + 4\pi a \int_{\mathbb{R}^3} |u(x)|^4 \, dx \tag{1-8}$$

in a suitable sense. Note that the occurrence of the scattering length in (1-8) is subtle: this functional is *not* obtained by testing H_N with factorized states of the form $u^{\otimes N}$ (which would lead to a functional with $4\pi a$ replaced by $\frac{1}{2}\int w$). Taking into account the short-range correlation structure which gives rise to (1-8) is the main difficulty in the proof of the following theorem, which is our main result.

Theorem 1.1 (Derivation of the Gross–Pitaevskii functional).
Under conditions (1-2), (1-3) *and* (1-4), *we have*

$$\lim_{N\to\infty} \inf_{\|\Psi\|_{\mathfrak{H}^N}=1} \frac{\langle \Psi, H_N \Psi\rangle}{N} = \inf_{\|u\|_{L^2(\mathbb{R}^3)}=1} \mathcal{E}_{\mathrm{GP}}(u) =: e_{\mathrm{GP}}. \tag{1-9}$$

Moreover, if Ψ_N *is an approximate ground state for* H_N, *namely*

$$\lim_{N\to\infty} \frac{\langle \Psi_N, H_N \Psi_N\rangle}{N} = e_{\mathrm{GP}},$$

then there exists a subsequence Ψ_{N_ℓ} *and a Borel probability measure* μ *supported on the set of minimizers of* $\mathcal{E}_{\mathrm{GP}}(u)$ *such that*

$$\lim_{\ell\to\infty} \operatorname{Tr}\left| \gamma^{(k)}_{\Psi_{N_\ell}} - \int |u^{\otimes k}\rangle\langle u^{\otimes k}| \, d\mu(u) \right| = 0 \quad \forall k \in \mathbb{N}, \tag{1-10}$$

where $\gamma^{(k)}_{\Psi_N} = \mathrm{Tr}_{k+1\to N}\,|\Psi_N\rangle\langle\Psi_N|$ *is the* k*-particle reduced density matrix of* Ψ_N*. In particular, if* $\mathcal{E}_{\rm GP}(u)$ *subject to* $\|u\|_{L^2}=1$ *has a unique minimizer* u_0 *(up to a complex phase), then there is complete Bose–Einstein condensation*

$$\lim_{N\to\infty}\mathrm{Tr}\big|\gamma^{(k)}_{\Psi_N}-|u_0^{\otimes k}\rangle\langle u_0^{\otimes k}|\big|=0\quad\forall k\in\mathbb{N}.\tag{1-11}$$

The energy upper bound in (1-9) was proved in [Lieb et al. 2000; Seiringer 2003] (see also [Benedikter et al. 2016, Appendix A] for an alternative approach). The energy lower bound in (1-9) and the convergence of one-particle density matrices were proved in [Lieb and Seiringer 2006]. The simpler case $A\equiv 0$ has been treated before in [Lieb et al. 2000] (ground-state energy) and [Lieb and Seiringer 2002] (condensation). In this case, the uniqueness of the Gross–Pitaevskii minimizer u_0 follows from a simple convexity argument. The result in Theorem 1.1 is thus not new, but the existing proofs are fairly difficult, in particular that of [Lieb and Seiringer 2006] which deals with the case $A\not\equiv 0$.

In the present paper, we will provide alternative proofs of the energy lower bound and the convergence of states using the quantum de Finetti theorem in the same spirit as in [Lewin et al. 2014; 2015a]. Our proofs are conceptually and technically simpler than those provided in [Lieb and Seiringer 2006]. The overall strategy will be explained in the next section.

Our result covers the case of a rotating gas where the minimizers of the Gross–Pitaevskii functional can develop quantized vortices. This corresponds to taking $A(x)=\Omega\wedge x$, with Ω being the angular velocity vector. In this case, V should be interpreted as the trapping potential minus $\frac{1}{2}(\Omega\wedge x)^2$. The assumption $V(x)\to\infty$ as $|x|\to\infty$ is to ensure that all particles are confined to the system. Here our conditions on A and V are slightly more general than those of [Lieb and Seiringer 2006], where A is assumed to grow at most polynomially and V is assumed to grow at least logarithmically.

The finite range assumption on w is not a serious restriction because we can always restrict the support of w to a finite ball without changing the scattering length significantly. In fact, it is sufficient to assume that w is integrable at infinity, in which case the scattering length is well-defined. We can also work with a more general interaction $w_N\geqslant 0$ (with scattering length a_N) rather than the specific choice (1-4), as long as its range goes to zero and $\lim_{N\to\infty}Na_N$ exists; then the result in Theorem 1.1 still holds with a replaced by $\lim_{N\to\infty}Na_N$. In particular, if w_N is chosen as in (1-6) for some $0<\beta<1$, then $Na_N\to(8\pi)^{-1}\int w$. The critical case $\beta=1$ considered in this paper is much more interesting because in the limit, the true scattering length appears instead of its first-order Born approximation $(8\pi)^{-1}\int w$.

2. Overall strategy

In this section we give an outline of the proofs of our main results, in order to better emphasize the key new points for the energy lower bound and the convergence of states.

We shall use the following notation: Let $\theta:\mathbb{R}^3\to\mathbb{R}$ be a radial smooth Heaviside-like function; i.e.,

$$0\leqslant\theta\leqslant 1,\qquad \theta(x)\equiv 0\quad\text{for } |x|\leqslant 1\qquad\text{and}\qquad \theta(x)\equiv 1\quad\text{for } |x|\geqslant 2.$$

Let $U : \mathbb{R}^3 \to \mathbb{R}$ be a radial smooth function supported on the annulus $\frac{1}{2} \leqslant |x| \leqslant 1$ such that

$$U(x) \geqslant 0 \quad \text{and} \quad \int_{\mathbb{R}^3} U = 4\pi a.$$

For every $R > 0$, define

$$\theta_R(x) = \theta\Big(\frac{x}{R}\Big), \quad U_R(x) = \frac{1}{R^3} U\Big(\frac{x}{R}\Big).$$

The smooth cut-off function θ_R will be used to perform cut-offs in both space and momentum variables, the latter being always denoted by

$$p = -i\nabla.$$

The potential U_R will be used to replace the original one. The important points will be that the integral of U_R yields the correct physical scattering length, and that we will have some freedom in choosing the range R of U_R.

Step 1 (Dyson's lemma). The main difficulty in dealing with the Gross–Pitaevskii limit is that an ansatz $u^{\otimes N}$ does *not* give the correct energy asymptotics. In this regime, correlations between particles *do* matter, and one should rather think of an ansatz of the form

$$\prod_{i=1}^{N} u(x_i) \prod_{1 \leqslant i < j \leqslant N} f(x_i - x_i), \tag{2-1}$$

or a close variant, where f is linked to the two-body scattering process. We shall follow the approach of [Lieb and Seiringer 2006], relying on a generalization of an idea due to Dyson [1957]. The following lemma, proved in [Lieb et al. 2005a], allows us to bound our Hamiltonian from below by an effective one which is much less singular, but still encodes the scattering length of the original interaction potential.

Lemma 2.1 (Generalized Dyson lemma).
For all $s > 0$, $1 > \varepsilon > 0$ and $R > 2R_0/N$, we have

$$H_N \geqslant \sum_{j=1}^{N} \big(h_j - (1-\varepsilon) p_j^2 \theta_s(p_j)\big) + \frac{(1-\varepsilon)^2}{N} W_N - C\frac{N^2 R^2 s^5}{\varepsilon}, \tag{2-2}$$

where

$$W_N := \sum_{i \neq j}^{N} U_R(x_i - x_j) \prod_{k \neq i,j} \theta_{2R}(x_j - x_k). \tag{2-3}$$

Here and in the sequel, C stands for a generic positive constant.

Proof. Recall that the scattering length of w_N is a/N. Therefore, from equation (50) and the first estimate in (52) in [Lieb et al. 2005a], with (v, a, χ, s) replaced by $(w_N, a/N, \theta_s, s^{-1})$, one has

$$p^2 \theta_s(p) + \frac{1}{2} \sum_{j=1}^{N-1} w_N(x - y_j) \geqslant \frac{1-\varepsilon}{N} \sum_{j=1}^{N-1} U_R(x - y_j) - \frac{CaR^2 s^5}{\varepsilon}$$

on $L^2(\mathbb{R}^3)$ for all given points y_j satisfying $\min_{j \neq k} |y_j - y_k| \geqslant 2R$. Since the left side is nonnegative, we can relax the condition $\min_{j \neq k} |y_j - y_k| \geqslant 2R$ by multiplying the right side by $\prod_{k \neq j} \theta_{2R}(y_j - y_k)$.

Thus for every $i = 1, 2, \dots, N$,

$$p_i^2\theta_s(p_i) + \frac{1}{2}\sum_{j\neq i}^{N} w_N(x_i - x_j) \geqslant \frac{1-\varepsilon}{N}\sum_{j\neq i} U_R(x_i - x_j)\prod_{k\neq i,j}\theta_{2R}(x_j - x_k) - \frac{CaR^2s^5}{\varepsilon}.$$

Multiplying both sides by $1-\varepsilon$ and summing over i, we obtain (2-2). □

Clarification. The reader should keep in mind that we will choose $R = R(N) \to 0$ (actually $N^{-1/2} \gg R \gg N^{-2/3}$), then $s \to \infty$ and $\varepsilon \to 0$.

The main point of Dyson's lemma is that we can replace the hard interaction potential w_N by a softer one U_R which encodes the scattering length conveniently as $\int U_R = 4\pi a$. The price we have to pay for this advantage is twofold, however: first, we have to use all the high-momentum part of the kinetic energy (note that $\theta_s(p) = 1$ when $p \geqslant 2s$); and second, the new potential $U_R(x_i - x_j)$ comes with the cut-off $\prod_{k\neq i,j}\theta(x_j - x_k)$. Together they really describe a "nearest neighbor" potential instead of an ordinary two-body potential. While the first problem is not too annoying, as the low part of the momentum is sufficient to recover the full energy in the limit, the second problem is much more serious.

Step 2 (Second moment estimate). The lower bound (2-2) leads us to consider the effective Hamiltonian

$$\widetilde{H}_N := \sum_{j=1}^{N}\tilde{h}_j + \frac{(1-\varepsilon)^2}{N}W_N, \tag{2-4}$$

where

$$\tilde{h} := h - (1-\varepsilon)p^2\theta_s(p) - \kappa_{\varepsilon,s}, \quad \kappa_{\varepsilon,s} := \inf\sigma\big(h - (1-\varepsilon)p^2\theta_s(p) - 1\big). \tag{2-5}$$

Here we use the freedom to add and remove the constant $N\kappa_{\varepsilon,s}$ to the Hamiltonian to reduce to the case $\tilde{h} \geqslant 1$. In order to ensure that $\kappa_{\varepsilon,s}$ is finite, we need the extra condition

$$\lim_{|x|\to\infty}\frac{|A(x)|^2}{V(x)} = 0, \tag{2-6}$$

which can be removed at a later stage, as we shall explain below.

We will now seek a lower bound to the ground-state energy of (2-4). The philosophy, as in the previous work [Lieb and Seiringer 2006], is that if Ψ_N is the ground state of the original Hamiltonian, then roughly

$$\Psi_N \approx \widetilde{\Psi}_N \prod_{1\leqslant i<j\leqslant N} f(x_i - x_i),$$

where f encodes the two-body scattering process and $\widetilde{\Psi}_N$ is a ground state for (2-4). Thus the Dyson lemma allows to extract the short-range correlation structure, and we now want to justify that $\widetilde{\Psi}_N$ can be approximated by a tensor power $u^{\otimes N}$; that is, we want to justify the mean-field approximation at the level of the ground state of (2-4).

There are two key difficulties left:

- The effective Hamiltonian is genuinely many-body. It can be bounded below by a three-body Hamiltonian, but obviously one will ultimately have to show that the three-body contribution can be neglected.

- To recover the correct energy in the limit, we need to take $R \ll N^{-1/3}$ in order to be able to neglect the three-body contribution in the effective Hamiltonian. We thus still have to deal with the mean-field approximation in the "rare but strong collisions" limit. In other words, even though the effective Hamiltonian is much less singular than the original one, we do not have the freedom to reduce the singularity as much as we would like.

It is in treating these two difficulties that our new method significantly departs from the previous works [Lewin et al. 2015a; Lieb and Seiringer 2006]. We shall rely on a strong a priori estimate for ground states of (2-4). In Lemma 3.1, we assume (2-6) and show that (provided $R \gg N^{-2/3}$, which is sufficient for our purpose)

$$(\widetilde{H}_N)^2 \geqslant \frac{1}{3}\Bigl(\sum_{j=1}^{N} \tilde{h}_j\Bigr)^2. \tag{2-7}$$

Note that a bound of this kind is not available for the original H_N due to the singularity of its interaction potential. In particular, (2-7) implies that every ground state $\widetilde{\Psi}_N$ of $\widetilde{H}_N$ satisfies the strong a priori estimate

$$\langle \widetilde{\Psi}_N, \tilde{h}_1 \tilde{h}_2 \widetilde{\Psi}_N \rangle \leqslant C_{\varepsilon,s}. \tag{2-8}$$

This second moment estimate is the key point in our analysis in the next steps. It is reminiscent of similar estimates used in the literature for the time-dependent problem [Erdős et al. 2007; 2009; 2010; Erdős and Yau 2001].

Notation. We always denote by C_ε (or $C_{\varepsilon,s}$) a (generic) constant independent of s, N and R (or independent of N and R, respectively).

Step 3 (Three-body estimate). Next we have to remove the cut-off

$$\prod_{k \neq i,j} \theta(x_j - x_k)$$

in W_N to obtain a lower bound in terms of a two-body Hamiltonian. Using the elementary inequality (see [Lieb and Seiringer 2006, equation (22)])

$$\prod_{k:k\neq i,j} \theta_{2R}(x_j - x_k) \geqslant 1 - \sum_{k:k\neq i,j} (1 - \theta_{2R}(x_j - x_k)),$$

we have

$$W_N \geqslant \sum_{i\neq j}^{N} U_R(x_i - x_j) - \sum_{k\neq i\neq j\neq k} U_R(x_i - x_j)(1 - \theta_{2R}(x_j - x_k)), \tag{2-9}$$

and we thus have only a three-body term to estimate. Since the summand in this term is zero except when $|x_i - x_j| \leqslant R$ and $|x_j - x_k| \leqslant 4R$, the last sum of (2-9) can be removed if the probability of having three or more particles in a region of diameter $O(R)$ is small enough. This should be the case if R is much smaller than $N^{-1/3}$, the average distance between particles, but it is rather difficult to confirm this intuition rigorously.

In [Lieb and Seiringer 2006], a three-body estimate was established using a subtle argument based on path integrals (the Trotter product formula). In this paper, we will follow a different, simpler approach. Instead of working directly with a ground state of H_N as in [Lieb and Seiringer 2006], we will consider a ground state $\widetilde{\Psi}_N$ of the effective Hamiltonian $\widetilde{H}_N$. Thanks to the second moment estimate (2-7), we can show that (see Lemma 3.4)

$$\sum_{k=3}^{N} \big\langle \widetilde{\Psi}_N, U_R(x_1-x_2)\theta_{2R}(x_2-x_k)\widetilde{\Psi}_N \big\rangle \leqslant C_{\varepsilon,s} N R^2. \tag{2-10}$$

The right side of (2-10) is small with our choice $N^{-1/2} \gg R$.

Step 4 (Mean-field approximation). With the cut-off in W_N removed, $\widetilde{H}_N$ turns into the two-body Hamiltonian

$$K_N := \sum_{j=1}^{N} \tilde{h}_j + \frac{(1-\varepsilon)^2}{N} \sum_{i \neq j} U_R(x_i - x_j)$$

for which we can validate the *mean-field approximation*. This is the simplest approximation for Bose gases where one restricts the many-body wave functions to the pure tensor products $u^{\otimes N}$. Since U_R converges to the delta-interaction with mass $\int U_R = 4\pi a$, we formally obtain the following approximation for the ground-state energy

$$e_{\mathrm{NL}}(\varepsilon, s) := \inf_{\|u\|_{L^2}=1} \left(\langle u, \tilde{h} u \rangle + (1-\varepsilon)^2 4\pi a \int |u|^4 \right).$$

In Section 4A, we will show that

$$\lim_{N\to\infty} \frac{\inf \sigma(K_N)}{N} = e_{\mathrm{NL}}(\varepsilon, s). \tag{2-11}$$

A similar result was proved in [Lieb and Seiringer 2006] using a coherent state method, which is a generalization of the c-number substitution in [Lieb et al. 2005c]. In the present paper, we will provide an alternative proof of (2-11) using the quantum de Finetti theorem of Størmer [1969] and Hudson and Moody [1976]. We note that this theorem has proved useful also in the derivation of the Gross–Pitaevskii equation in the dynamical case; see [Chen et al. 2015]. The following formulation is taken from [Lewin et al. 2014, Corollary 2.4] (see [Rougerie 2015] for a general discussion and more references):

Theorem 2.2 (Quantum de Finetti).
Let $\mathfrak{K}$ be an arbitrary separable Hilbert space and let $\Psi_N \in \bigotimes_{\mathrm{sym}}^N \mathfrak{K}$ with $\|\Psi_N\| = 1$. Assume that the sequence of one-particle density matrices $\gamma_{\Psi_N}^{(1)}$ converges strongly in trace class when $N \to \infty$. Then, up to a subsequence, there exists a (unique) Borel probability measure μ on the unit sphere $S\mathfrak{K}$, invariant under the group action of S^1, such that

$$\lim_{N\to\infty} \mathrm{Tr} \left| \gamma_{\Psi_N}^{(k)} - \int |u^{\otimes k}\rangle\langle u^{\otimes k}| \, d\mu(u) \right| = 0 \quad \forall k \in \mathbb{N}. \tag{2-12}$$

This theorem validates the mean-field approximation for a large class of trapped Bose gases, in particular (see [Lewin et al. 2014] and references therein) when the strength of the interaction is proportional to the inverse of the particle number, case $\beta = 0$ in (1-6). However, when the interaction becomes stronger, the mean-field approximation is harder to justify. The convergence (2-11) with $R \gg N^{-2/15}$ was proved in [Lewin et al. 2015a] by using a quantitative version of Theorem 2.2 valid for finite-dimensional spaces [Christandl et al. 2007; Chiribella 2011; Lewin et al. 2015b]. However, this range of R is too small for our purpose because we are forced to choose $R \ll N^{-1/2}$ in the previous steps.

In this paper, thanks to the strong a priori estimate (2-8), we are able to prove (2-11) for the larger range $R \gg N^{-2/3}$. As in [Lewin et al. 2015a; Lieb and Seiringer 2006], we localize the problem onto energy levels of the one-body Hamiltonian $\tilde{h}$ lying below a chosen cut-off Λ. At fixed Λ, it turns out that the projected Hamiltonian is bounded proportionally to N. We are thus in a usual mean-field scaling if we are allowed to take $N \to \infty$ first, and then $\Lambda \to \infty$ later. Taking limits in this order demands a very strong control on the localization error made by projecting the Hamiltonian, however. This control is provided again by the moment estimate (2-8).

Combining the arguments in Steps 1–4, we can pass to the limit $N \to \infty$; then $s \to \infty$ and $\varepsilon \to 0$ to obtain the energy convergence (1-9) under the extra condition (2-6). In Section 4B, we remove this technical assumption using a concavity argument from [Lieb and Seiringer 2006] and a binding inequality which goes back to an idea in [Lieb 1984].

Step 5 (Convergence of ground states). In Section 4C, we prove the convergence of (approximate) ground states using the convergence of the ground state energy of a perturbed Hamiltonian and the Feynman–Hellmann principle. A similar approach was used in [Lieb and Seiringer 2006] to prove the convergence of the 1-particle density matrix. However, the quantum de Finetti theorem helps us to avoid the complicated convex analysis in [Lieb and Seiringer 2006], simplifying the proof significantly and giving access to higher-order density matrices.

3. Second moment estimate

In this section, we consider the effective Hamiltonian obtained after applying the generalized Dyson lemma to the original problem, namely

$$\widetilde{H}_N = \sum_{j=1}^{N} \tilde{h}_j + \frac{(1-\varepsilon)^2}{N} W_N,$$

where $\tilde{h}$ and W_N are defined in (2-5) and (2-3), respectively. We will work under the extra assumption (2-6). Since $A \in L^3_{\mathrm{loc}}(\mathbb{R}^3, \mathbb{R}^3)$ and V grows faster than $|A|^2$ at infinity, for every $\varepsilon > 0$, we have

$$\left(\frac{V}{2} - 2\varepsilon^{-1}|A|^2\right)_- \in L^{3/2}(\mathbb{R}^3),$$

and hence

$$\left(\frac{\varepsilon}{4}\right)p^2 + \frac{V}{2} - 2\varepsilon^{-1}|A|^2 \geqslant -C_\varepsilon.$$

In combination with the Cauchy–Schwarz inequality, we get

$$h - (1-\varepsilon)p^2\theta_s(p) \geqslant \frac{\varepsilon}{2}p^2 - 2\varepsilon^{-1}|A|^2 + V \geqslant \frac{\varepsilon}{4}p^2 + \frac{V}{2} - C_\varepsilon.$$

Therefore, $\inf \sigma(h) - 1 \geqslant \kappa_{\varepsilon,s} \geqslant -C_\varepsilon$ and

$$\tilde{h} \geqslant C_\varepsilon(-\Delta + V + 1). \tag{3-1}$$

The key estimate in this section is the following:

Lemma 3.1 (Second moment estimate).
Assume that (2-6) *holds. For every* $1 > \varepsilon > 0$ *and* $s > 0$, *if*

$$R = R(N) \gg N^{-2/3}$$

when $N \to \infty$, *then for* N *large enough, we have the operator bound*

$$(\widetilde{H}_N)^2 \geqslant \frac{1}{3}\left(\sum_{j=1}^N \tilde{h}_j\right)^2. \tag{3-2}$$

We will show in Section 3C that a convenient lower bound to Dyson's potential W_N in terms of truly two-body operators follows from Lemma 3.1.

Before proving Lemma 3.1 in Section 3B, we first collect some useful inequalities on a generic translation-invariant interaction operator $W(x-y)$ that will be used throughout the paper.

3A. ***Operator inequalities for interaction potentials.*** We state several useful inequalities in the following lemma. In fact, (3-3) is well-known, and (3-4) with $\delta = 0$ was proved earlier in [Erdős and Yau 2001, Lemma 5.3]. In the sequel, we will crucially rely on the improvement to $\delta > 0$, and on (3-5), which seem to be new.

Lemma 3.2 (Inequalities for a repulsive interaction potential).
For every $0 \leqslant W \in L^1 \cap L^2(\mathbb{R}^3)$, *the multiplication operator* $W(x-y)$ *on* $L^2((\mathbb{R}^3)^2)$ *satisfies*

$$0 \leqslant W(x-y) \leqslant C\|W\|_{L^{3/2}(\mathbb{R}^3)}(-\Delta_x), \tag{3-3}$$

and, for any $0 \leqslant \delta < \frac{1}{4}$,

$$0 \leqslant W(x-y) \leqslant C_\delta \|W\|_{L^1(\mathbb{R}^3)}(1-\Delta_x)^{1-\delta}(1-\Delta_y)^{1-\delta}. \tag{3-4}$$

Moreover, for all $1 > \varepsilon > 0$, $s > 0$, $A \in L^3_{\text{loc}}(\mathbb{R}^3, \mathbb{R}^3)$ *and* $0 \leqslant V \in L^1_{\text{loc}}(\mathbb{R}^3)$,

$$\tilde{h}_x W(x-y) + W(x-y)\tilde{h}_x \geqslant -C\big(\|W\|_{L^2} + (1+s^2)\|W\|_{L^{3/2}}\big)(1-\Delta_x)(1-\Delta_y). \tag{3-5}$$

Proof of Lemma 3.2. We will prove this in several steps.

Proof of (3-3). From Hölder's and Sobolev's inequalities, we have

$$\begin{aligned}\langle f, W(x-y)f\rangle &= \iint W(x-y)|f(x,y)|^2\,dx\,dy \\ &\leqslant \int\left(\int W(x-y)^{3/2}\,dx\right)^{2/3}\left(\int |f(x,y)|^6\,dx\right)^{1/3} dy \\ &\leqslant C\|W\|_{L^{3/2}(\mathbb{R}^3)}\int\left(\int |\nabla_x f(x,y)|^2\,dx\right)dy\end{aligned}$$

for every function $f \in H^1((\mathbb{R}^3)^2)$. Therefore, (3-3) follows immediately.

Proof of (3-4). The estimate (3-4) with $\delta = 0$ was first derived in [Erdős and Yau 2001]. The following proof is adapted from the proof (again for $\delta = 0$) in [Lieb and Seiringer 2006]. Note that for every operator K, we have $K^*K \leqslant 1$ if and only if $KK^* \leqslant 1$. Therefore, (3-4) is equivalent to

$$\sqrt{W(x-y)}(1-\Delta_x)^{\delta-1}(1-\Delta_y)^{\delta-1}\sqrt{W(x-y)} \leqslant C_\delta \|W\|_{L^1}. \tag{3-6}$$

Let G be the Green function of $(1-\Delta)^{\delta-1}$ whose Fourier transform is given by

$$\widehat{G}(k) := \int_{\mathbb{R}^3} e^{-2\pi i x\cdot k} G(x)\,dx = \frac{1}{(1+4\pi^2|k|^2)^{1-\delta}}.$$

For every function $f \in L^2((\mathbb{R}^3)^2)$, one has

$$\begin{aligned}
&\langle f, \sqrt{W(x-y)}(1-\Delta_x)^{\delta-1}(1-\Delta_y)^{\delta-1}\sqrt{W(x-y)}\, f\rangle \\
&\qquad = \int \overline{f(x,y)}\sqrt{W(x-y)}G(x-x')G(y-y')\sqrt{W(x'-y')}\, f(x',y')\,dx\,dy\,dx'\,dy' \\
&\qquad \leqslant \int \frac{W(x-y)|G(x-x')|^2|f(x',y')|^2 + W(x'-y')|G(y-y')|^2|f(x,y)|^2}{2} \\
&\qquad = C_\delta \|W\|_{L^1}\langle f, f\rangle,
\end{aligned}$$

where

$$C_\delta := \int |G|^2 = \int |\widehat{G}|^2 = \int_{\mathbb{R}^3} \frac{dk}{(1+4\pi^2|k|^2)^{2(1-\delta)}},$$

which is finite for all $0 \leqslant \delta < \frac{1}{4}$. Thus (3-6), and hence (3-4), holds true.

Simpler version of (3-5). We are going to deduce (3-5) from the inequality

$$(-\Delta_x)W(x-y) + W(x-y)(-\Delta_x) \geqslant -C\big(\|W\|_{L^{3/2}} + \|W\|_{L^2}\big)(1-\Delta_x)(1-\Delta_y). \tag{3-7}$$

By an approximation argument, one can assume that W is smooth. For every $f \in H^2(\mathbb{R}^3\times\mathbb{R}^3)$, a straightforward calculation using integration by parts, and the identity $\nabla_x(W(x-y)) = -\nabla_y(W(x-y))$ gives us

$$\begin{aligned}
&\big\langle f, \big((-\Delta_x)W(x-y) + W(x-y)(-\Delta_x)\big) f\big\rangle \\
&\qquad = 2\Re \iint \nabla_x \overline{f(x,y)}\nabla_x(W(x-y)f(x,y))\,dx\,dy \\
&\qquad = 2\iint |\nabla_x f(x,y)|^2 W(x-y) + 2\Re\iint \nabla_x\overline{f(x,y)}\nabla_x(W(x-y))f(x,y)\,dx\,dy \\
&\qquad \geqslant -2\Re\iint \nabla_x\overline{f(x,y)}\nabla_y(W(x-y))f(x,y)\,dx\,dy \\
&\qquad = 2\Re\iint \nabla_y\big((\nabla_x\overline{f(x,y)})f(x,y)\big)W(x-y)\,dx\,dy \\
&\qquad = 2\Re\iint \big(\nabla_x\overline{f(x,y)}\nabla_y f(x,y) + \nabla_y(\nabla_x\overline{f(x,y)})f(x,y)\big)W(x-y)\,dx\,dy.
\end{aligned}$$

Using Cauchy–Schwarz and Sobolev's inequality (3-3), we get

$$\left|\iint \nabla_x \overline{f(x,y)} \nabla_y f(x,y) W(x-y)\, dx\, dy\right| \leqslant \iint \frac{|\nabla_x f(x,y)|^2 + |\nabla_y f(x,y)|^2}{2} |W(x-y)|\, dx\, dy$$
$$\leqslant C\|W\|_{L^{3/2}} \langle f, (-\Delta_x)(-\Delta_y) f\rangle.$$

Moreover, by the Cauchy–Schwarz inequality again and (3-4) (with $\delta = 0$ and W replaced by W^2),

$$\left|\iint \big(\nabla_y \nabla_x \overline{f(x,y)}\big) f(x,y) W(x-y)\, dx\, dy\right|$$
$$\leqslant \left(\iint |\nabla_y \nabla_x f(x,y)|^2\, dx\, dy\right)^{1/2} \left(\iint |f(x,y)|^2 |W(x-y)|^2\, dx\, dy\right)^{1/2}$$
$$\leqslant C\|W\|_{L^2} \langle f, (1-\Delta_x)(1-\Delta_y) f\rangle.$$

Thus we obtain

$$\big\langle f, \big((-\Delta_x) W(x-y) + W(x-y)(-\Delta_x)\big) f\big\rangle \geqslant -C(\|W\|_{L^{3/2}} + \|W\|_{L^2}) \langle f, (1-\Delta_x)(1-\Delta_y) f\rangle$$

for all $f \in H^2(\mathbb{R}^3 \times \mathbb{R}^3)$. This proves (3-7).

Proof of (3-5). From the commutator relation

$$p_x W(x-y) = W(x-y) p_x + (-i\nabla_x W)(x-y),$$

we find that

$$\big(p_x A(x) + A(x) p_x + |A(x)|^2\big) W(x-y) + W(x-y)\big(p_x A(x) + A(x) p_x + |A(x)|^2\big)$$
$$= 2\big(p_x W(x-y) A(x) + A(x) W(x-y) p_x + |A(x)|^2 W(x-y)\big)$$
$$= 2(p_x + A(x)) W(x-y)(p_x + A(x)) - 2 p_x W(x-y) p_x.$$

Using

$$(p_x + A(x)) W(x-y)(p_x + A(x)) \geqslant 0$$

and estimating $p_x W(x-y) p_x$ by Sobolev's inequality (3-3), we get

$$\big(p_x A(x) + A(x) p_x + |A(x)|^2\big) W(x-y) + W(x-y)\big(p_x A(x) + A(x) p_x + |A(x)|^2\big)$$
$$\geqslant -C\|W\|_{L^{3/2}} (-\Delta_x)(-\Delta_y). \tag{3-8}$$

Finally, by (3-3) again and the Cauchy–Schwarz inequality for operators

$$\pm(XY + Y^*X^*) \leqslant \delta XX^* + \delta^{-1} Y^*Y \quad \forall \delta > 0, \tag{3-9}$$

we obtain

$$p_x^2(1-\theta_s(p_x)) W(x-y) + W(x-y) p_x^2(1-\theta_s(p_x))$$
$$\geqslant -\delta p_x^2(1-\theta_s(p_x)) W(x-y) p_x^2(1-\theta_s(p_x)) + \delta^{-1} W(x-y)$$
$$\geqslant -C\|W\|_{L^{3/2}} \big(\delta p_x^4 (1-\theta_s(p_x))^2 + \delta^{-1}\big)(-\Delta_x)$$

for all $\delta > 0$. Using $1 - \theta_s(p) \leqslant \mathbb{1}(|p| \leqslant 2s)$ and choosing $\delta \sim s^{-2}$ gives

$$p_x^2(1-\theta_s(p_x)) W(x-y) + W(x-y) p_x^2(1-\theta_s(p_x)) \geqslant -Cs^2 \|W\|_{L^{3/2}} (-\Delta_x). \tag{3-10}$$

From (3-7), (3-8) and (3-10), the bound (3-5) follows. □

3B. ***Proof of Lemma 3.1.*** Before completing the proof of Lemma 3.1, we make a remark on the simpler case with the Dyson potential W_N replaced by a truly two-body interaction.

Remark 3.3 (Second moment estimate with two-body interactions).
Consider the model case

$$K_N := \sum_{j=1}^{N} \tilde{h}_j + \frac{(1-\varepsilon)^2}{N} \sum_{i \neq j} U_R(x_i - x_j).$$

By expanding K_N^2 and using the fact that $\tilde{h}_i \geqslant 0$ commutes with $U_R(x_j - x_k) \geqslant 0$ when $i \neq j$ and $i \neq k$, and then using (3-5) to estimate terms of the form

$$\tilde{h}_i U_R(x_i - x_j) + U(x_i - x_j)\tilde{h}_i,$$

we obtain

$$K_N^2 \geqslant \frac{1}{3} \sum_{1 \leqslant i \neq j \leqslant N} \tilde{h}_i \tilde{h}_j \tag{3-11}$$

provided that $R = R(N) \gg N^{-2/3}$. A similar estimate also holds when $\tilde{h}$ is replaced by the original kinetic operator h.

We stress once again that we do *not* expect (3-11) to hold for our original Hamiltonian H_N, which is in the more singular regime $R \sim N^{-1}$. We thus need to work with the Dyson Hamiltonian, and its rather intricate nature makes the actual proof of Lemma 3.1 more difficult than the one we have sketched for (3-11). We now proceed with this proof.

Proof of Lemma 3.1. We have

$$(\widetilde{H}_N)^2 - \left(\sum_{j=1}^{N} \tilde{h}_j \right)^2 = \frac{(1-\varepsilon)^2}{N} \sum_{\ell=1}^{N} (\tilde{h}_\ell W_N + W_N \tilde{h}_\ell) + \frac{(1-\varepsilon)^4}{N^2} W_N^2. \tag{3-12}$$

As in Remark 3.3, the goal is to bound $\tilde{h}_1 W_N + W_N \tilde{h}_1$ from below. We first decompose the interaction operator as

$$W_N = W_a + W_b,$$

where

$$W_a = \sum_{1 \in \{i,j\}} U_R(x_i - x_j) \prod_{k \neq i,j} \theta_{2R}(x_j - x_k),$$

$$W_b = \sum_{i,j \geqslant 2} U_R(x_i - x_j) \prod_{k \neq i,j} \theta_{2R}(x_j - x_k).$$

Estimate of W_a. By the Cauchy–Schwarz inequality (3-9), we get

$$\pm(\tilde{h}_1 W_a + W_a \tilde{h}_1) \leqslant N^{-1} \tilde{h}_1 W_a \tilde{h}_1 + N W_a. \tag{3-13}$$

Let us show that

$$W_a \leqslant \frac{C}{R^3}. \tag{3-14}$$

Indeed, for every given $(x_1, x_2, \dots, x_N) \in (\mathbb{R}^3)^N$, the product

$$U_R(x_1 - x_j) \prod_{k \neq 1, j} \theta_{2R}(x_j - x_k)$$

is bounded by $\|U_R\|_{L^\infty} \leqslant CR^{-3}$ and it is zero except in the case

$$|x_1 - x_j| < R < 2R < \min_{k \neq 1, j} |x_j - x_k|.$$

By the triangle inequality, the latter condition implies that

$$|x_1 - x_j| < R < \min_{k \neq 1, j} |x_1 - x_k|,$$

and it is satisfied by at most one index $j \neq 1$. Therefore,

$$\sum_{j \geqslant 2} U_R(x_1 - x_j) \prod_{k \neq 1, j} \theta_{2R}(x_j - x_k) \leqslant \frac{C}{R^3}.$$

Similarly, we have

$$\sum_{i \geqslant 2} U_R(x_i - x_1) \prod_{k \neq 1, i} \theta_{2R}(x_1 - x_k) \leqslant \frac{C}{R^3},$$

and hence (3-14) holds true. From (3-13) and (3-14), we obtain

$$\pm(\tilde{h}_1 W_a + W_a \tilde{h}_1) \leqslant \frac{C}{NR^3}(\tilde{h}_1)^2 + 2N \sum_{1 \in \{i,j\}} U_R(x_i - x_j) \prod_{k \neq i, j} \theta_{2R}(x_j - x_k). \tag{3-15}$$

Here we do not need to estimate the second term on the right side of (3-15) because this term is part of W_N, which will be controlled by W_N^2 in $\widetilde{H}_N^2$.

Estimate of W_b. We need a further decomposition

$$W_b = \sum_{i,j \geqslant 2} U_R(x_i - x_j) \prod_{k \neq i, j} \theta_{2R}(x_j - x_k) = W_c - W_d,$$

where

$$W_c := \sum_{i,j \geqslant 2} U_R(x_i - x_j) \prod_{k \neq 1, i, j} \theta_{2R}(x_j - x_k),$$

$$W_d := \sum_{i,j \geqslant 2} U_R(x_i - x_j) \big(1 - \theta_{2R}(x_j - x_1)\big) \prod_{k \neq 1, i, j} \theta_{2R}(x_j - x_k).$$

Note that

$$W_c \geqslant 0, \quad W_d \geqslant 0 \quad \text{and} \quad \tilde{h}_1 W_c = W_c \tilde{h}_1 \geqslant 0.$$

On the other hand, by the Cauchy–Schwarz inequality (3-9) again,

$$\pm(\tilde{h}_1 W_d + W_d \tilde{h}_1) \leqslant \delta \tilde{h}_1 W_d \tilde{h}_1 + \delta^{-1} W_d. \tag{3-16}$$

We have two different ways to bound W_d. First, by (3-3) and (3-1),

$$\big(1 - \theta_{2R}(x_j - x_1)\big) \leqslant C \|1 - \theta_{2R}\|_{L^{3/2}} (1 - \Delta_1) \leqslant C_\varepsilon R^2 \tilde{h}_1.$$

Since here $i, j \geqslant 2$, both sides of the latter estimate commute with

$$U_R(x_i - x_j) \prod_{k \neq 1,i,j} \theta_{2R}(x_j - x_k),$$

and we deduce that

$$\big(1-\theta_{2R}(x_j-x_1)\big)U_R(x_i-x_j) \prod_{k\neq 1,i,j} \theta_{2R}(x_j-x_k) \leqslant C_\varepsilon R^2 \tilde{h}_1 U_R(x_i-x_j) \prod_{k\neq 1,i,j} \theta_{2R}(x_j-x_k).$$

Taking the sum over $i, j \geqslant 2$, we obtain

$$W_d \leqslant C_\varepsilon R^2 \tilde{h}_1 W_c. \tag{3-17}$$

Second, let us show that

$$W_d \leqslant \frac{C}{R^3}. \tag{3-18}$$

Indeed, for every given $(x_1, x_2, \dots, x_N) \in (\mathbb{R}^3)^N$, the product

$$U_R(x_i-x_j)\big(1-\theta_{2R}(x_j-x_1)\big) \prod_{k\neq 1,i,j} \theta_{2R}(x_j-x_k)$$

is zero except in the case

$$|x_i - x_j| < R, \quad |x_j - x_1| < 4R, \quad \min_{k\neq 1,i,j} |x_j - x_k| > 2R. \tag{3-19}$$

By the triangle inequality, (3-19) implies that the ball $B(x_1, 5R)$ contains $B\big(x_i, \frac12 R\big)$, $B\big(x_j, \frac12 R\big)$, and the balls $B\big(x_i, \frac12 R\big)$, $B\big(x_j, \frac12 R\big)$ do not intersect with $B\big(x_k, \frac12 R\big)$ for all $k \neq 1, i, j$. Since $B(x_1, 5R)$ can contain only a finite number of disjoint balls of radius $\frac12 R$, we see that there are only a finite number of pairs (i, j) satisfying (3-19). Thus we can conclude that

$$W_d \leqslant C \|U_R\|_{L^\infty} \leqslant C R^{-3}.$$

From (3-16), (3-17) and (3-18), we obtain

$$\tilde{h}_1 W_b + W_b \tilde{h}_1 = \tilde{h}_1 W_d + W_d \tilde{h}_1 + 2\tilde{h}_1 W_c \geqslant -\frac{C\delta}{R^3}(\tilde{h}_1)^2 + \left(2 - \frac{C_\varepsilon R^2}{\delta}\right)\tilde{h}_1 W_c.$$

Choosing $\delta \sim R^2$, we get

$$\tilde{h}_1 W_b + W_b \tilde{h}_1 \geqslant -\frac{C_\varepsilon}{R}(\tilde{h}_1)^2. \tag{3-20}$$

Conclusion. From (3-15) and (3-20), we get

$$\tilde{h}_1 W_N + W_N \tilde{h}_1 \geqslant -\left(\frac{C}{NR^3} + \frac{C_\varepsilon}{R}\right)(\tilde{h}_1)^2 - 2N \sum_{1\in\{i,j\}} U_R(x_i-x_j) \prod_{k\neq i,j} \theta_{2R}(x_j-x_k).$$

Summing the similar estimates with 1 replaced by ℓ and using

$$\sum_{\ell=1}^{N} \sum_{\ell\in\{i,j\}} U_R(x_i-x_j) \prod_{k\neq i,j} \theta_{2R}(x_j-x_k) = 2W_N,$$

we find that

$$\sum_{\ell=1}^{N}(\tilde{h}_\ell W_N + W_N\tilde{h}_\ell) \geqslant -\Big(\frac{C}{NR^3}+\frac{C_\varepsilon}{R}\Big)\sum_{\ell=1}^{N}(\tilde{h}_\ell)^2 - 2NW_N.$$

Therefore, coming back to (3-12), we conclude that (completing a square in the last inequality)

$$\begin{aligned}(\widetilde{H}_N)^2 - \Big(\sum_{j=1}^{N}\tilde{h}_j\Big)^2 &= \frac{(1-\varepsilon)^2}{N}\sum_{\ell=1}^{N}\Big(\tilde{h}_\ell W_N + W_N\tilde{h}_\ell\Big) + \frac{(1-\varepsilon)^4}{N^2}W_N^2 \\ &\geqslant -\Big(\frac{C}{N^2R^3}+\frac{C_\varepsilon}{NR}\Big)\sum_{\ell=1}^{N}(\tilde{h}_\ell)^2 - 2(1-\varepsilon)^2 W_N + \frac{(1-\varepsilon)^4}{N^2}W_N^2 \\ &\geqslant -\Big(\frac{C}{N^2R^3}+\frac{C_\varepsilon}{NR}\Big)\sum_{\ell=1}^{N}(\tilde{h}_\ell)^2 - N^2.\end{aligned}$$

When $R \gg N^{-2/3}$, we have

$$\frac{C}{N^2R^3}+\frac{C_\varepsilon}{NR} \ll 1,$$

and hence

$$(\widetilde{H}_N)^2 \geqslant 2\sum_{1\leqslant i<j\leqslant N}\tilde{h}_i\tilde{h}_j + (1-o(1))\sum_{\ell=1}^{N}(\tilde{h}_\ell)^2 - N^2,$$

which yields the result, recalling that in our convention, $\tilde{h} \geqslant 1$. □

3C. *Three-body estimate.* A first consequence of the second moment estimate in Lemma 3.1 is that we can conveniently bound Dyson's Hamiltonian from below by a two-body Hamiltonian. This is done by first using a simple bound in terms of a three-body Hamiltonian, and then bounding the unwanted three-body part.

Lemma 3.4 (Three-body estimate).
Assume the extra condition (2-6) *holds. For every* $1 > \varepsilon > 0$ *and* $s > 0$, *if* $R = R(N) \gg N^{-2/3}$, *then*

$$\sum_{i\neq j}U_R(x_i-x_j)\sum_{k\neq i,j}(1-\theta_{2R}(x_j-x_k)) \leqslant C_{\varepsilon,s}\frac{R^2}{N}(\widetilde{H}_N)^4. \tag{3-21}$$

Consequently,

$$\widetilde{H}_N \geqslant \sum_{j=1}^{N}\tilde{h}_j + \frac{(1-\varepsilon)^2}{N}\sum_{i\neq j}U_R(x_i-x_j) - C_{\varepsilon,s}\frac{R^2}{N^2}(\widetilde{H}_N)^4. \tag{3-22}$$

Note the error term involving $(\widetilde{H}_N)^4$, which is well under control since we are interested in its expectation value in a ground state.

Proof. By (3-3) and (3-1), we have

$$(1-\theta_{2R}(x_2-x_k)) \leqslant C_{\varepsilon,s}R^2\tilde{h}_k \quad \text{for } k \geqslant 3.$$

Since $U_R(x_1-x_2)$ commutes with both sides, we get

$$
\begin{aligned}
U_R(x_1-x_2)\sum_{k\geqslant 3}(1-\theta_{2R}(x_2-x_k)) &\leqslant C_{\varepsilon,s}R^2 U_R(x_1-x_2)\sum_{k\geqslant 3}\tilde{h}_3\\
&= \tfrac{1}{2}C_{\varepsilon,s}R^2\big(\widetilde{H}_N-\tilde{h}_1-\tilde{h}_2-(1-\varepsilon)^2N^{-1}W_N\big)U_R(x_1-x_2)\\
&\quad +\tfrac{1}{2}C_{\varepsilon,s}R^2 U_R(x_1-x_2)\big(\widetilde{H}_N-\tilde{h}_1-\tilde{h}_2-(1-\varepsilon)^2N^{-1}W_N\big)\\
&\leqslant \tfrac{1}{2}C_{\varepsilon,s}R^2\big(\widetilde{H}_N U_R(x_1-x_2)+U_R(x_1-x_2)\widetilde{H}_N\big)\\
&\quad +\tfrac{1}{2}C_{\varepsilon,s}R^2\sum_{j=1}^{2}\big(\tilde{h}_j U_R(x_1-x_2)+U_R(x_1-x_2)\tilde{h}_j\big). \qquad (3\text{-}23)
\end{aligned}
$$

In the last estimate, we have used $W_N\geqslant 0$. Thanks to (3-5) and (3-1), we get for all $j=1,2$,

$$
\tilde{h}_j U_R(x_1-x_2)+U_R(x_1-x_2)h_j \geqslant -C_{\varepsilon,s}R^{-3/2}(1-\Delta_1)(1-\Delta_2)\geqslant -C_{\varepsilon,s}R^{-3/2}\tilde{h}_1\tilde{h}_2. \qquad (3\text{-}24)
$$

On the other hand, by the Cauchy–Schwarz inequality (3-9) and (3-4) (with $\delta=0$ and $W=U_R$),

$$
\begin{aligned}
\widetilde{H}_N U_R(x_1-x_2)+U_R(x_1-x_2)\widetilde{H}_N &\leqslant \delta\widetilde{H}_N U_R(x_1-x_2)\widetilde{H}_N+\delta^{-1}U_R(x_1-x_2)\\
&\leqslant C_{\varepsilon,s}\delta\widetilde{H}_N\tilde{h}_1\tilde{h}_2\widetilde{H}_N+C_{\varepsilon,s}\delta^{-1}\tilde{h}_1\tilde{h}_2 \qquad (3\text{-}25)
\end{aligned}
$$

for all $\delta>0$. Choosing $\delta=N^{-1}$ and using $R^{-3/2}\leqslant N$, we deduce from (3-23), (3-24) and (3-25) that

$$
U_R(x_1-x_2)\sum_{k\geqslant 3}(1-\theta_{2R}(x_2-x_k))\leqslant C_{\varepsilon,s}R^2\big(N^{-1}\widetilde{H}_N\tilde{h}_1\tilde{h}_2\widetilde{H}_N+N\tilde{h}_1\tilde{h}_2\big).
$$

By symmetrization with respect to the indices, we find that

$$
\sum_{i\neq j}U_R(x_1-x_2)\sum_{k\neq i,j}(1-\theta_{2R}(x_j-x_k))\leqslant C_{\varepsilon,s}R^2\bigg(N^{-1}\widetilde{H}_N\sum_{i\neq j}\tilde{h}_i\tilde{h}_j\widetilde{H}_N+N\sum_{i\neq j}\tilde{h}_i\tilde{h}_j\bigg).
$$

Combining with the second moment estimate (3-2), we obtain (3-21). From the three-body estimate (3-21) and the elementary inequality (2-9), the operator bound (3-22) follows. □

4. Energy lower bound and convergence of states

4A. *Mean-field approximation.* We are now reduced to justifying the mean-field approximation for a new Hamiltonian with the two-body interaction $U_R(x-y)$, which converges to a Dirac delta much slower than the original one. The analysis in this section provides an alternative to the coherent states method of [Lieb and Seiringer 2006].

Proposition 4.1 (Mean-field approximation).
Assume that (2-6) *holds. For every* $1>\varepsilon>0$ *and* $s>0$, *if*

$$
N^{-1/2}\gg R=R(N)\gg N^{-2/3}
$$

then

$$\lim_{N\to\infty}\frac{\inf\sigma(\widetilde{H}_N)}{N}=\inf_{\|u\|_{L^2}=1}\left(\langle u,\tilde{h}u\rangle+(1-\varepsilon)^2 4\pi a\int|u|^4\right)=:e_{\mathrm{NL}}(\varepsilon,s). \tag{4-1}$$

Proof. The upper bound in (4-1) can be obtained easily using trial states of the form $u^{\otimes N}$. For the lower bound, let us consider a ground state $\widetilde{\Psi}_N$ of $\widetilde{H}_N$ (which exists because $\tilde{h}$ has compact resolvent). Using the ground-state equation, we find that

$$\langle\widetilde{\Psi}_N,(\widetilde{H}_N)^k\widetilde{\Psi}_N\rangle=(\inf\sigma(\widetilde{H}_N))^k\leqslant(C_{\varepsilon,s}N)^k \tag{4-2}$$

for all $k\in\mathbb{N}$. In particular, the second moment estimate (3-2) implies that

$$\langle\widetilde{\Psi}_N,\tilde{h}_1\tilde{h}_2\widetilde{\Psi}_N\rangle\leqslant C_{\varepsilon,s}, \tag{4-3}$$

and the operator estimate (3-22) implies that

$$\liminf_{N\to\infty}\frac{\langle\widetilde{\Psi}_N,\widetilde{H}_N\widetilde{\Psi}_N\rangle}{N}\geqslant\liminf_{N\to\infty}\big(\mathrm{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N})+(1-\varepsilon)^2\,\mathrm{Tr}(U_R\gamma^{(2)}_{\widetilde{\Psi}_N})\big). \tag{4-4}$$

Here $\gamma^{(k)}_{\widetilde{\Psi}_N}$ is the k-particle density matrix of $\widetilde{\Psi}_N$ and U_R is understood as the multiplication operator $U_R(x-y)$ on $\mathfrak{H}^2$. Since $\mathrm{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N})$ is bounded uniformly in N and $\tilde{h}$ has compact resolvent, up to a subsequence we can assume that $\gamma^{(1)}_{\widetilde{\Psi}_N}$ converges strongly in trace class. By the quantum de Finetti theorem, Theorem 2.2, up to a subsequence we can find a Borel probability measure $\tilde{\mu}$ on the unit sphere $S\mathfrak{H}$ such that

$$\lim_{N\to\infty}\mathrm{Tr}\left|\gamma^{(k)}_{\widetilde{\Psi}_N}-\int|u^{\otimes k}\rangle\langle u^{\otimes k}|\,d\tilde{\mu}(u)\right|=0\quad\forall k\in\mathbb{N}. \tag{4-5}$$

We will show that

$$\liminf_{N\to\infty}\big(\mathrm{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N})+(1-\varepsilon)^2\,\mathrm{Tr}(U_R\gamma^{(2)}_{\widetilde{\Psi}_N})\big)\geqslant\int\left(\langle u,\tilde{h}u\rangle+(1-\varepsilon)^2 4\pi a\int|u|^4\right)d\tilde{\mu}(u), \tag{4-6}$$

and then the lower bound in (4-1) follows immediately. Since $\tilde{h}$ is positive and independent of N, (4-5) and Fatou's lemma imply

$$\liminf_{N\to\infty}\mathrm{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N})\geqslant\int\langle u,\tilde{h}u\rangle\,d\tilde{\mu}(u). \tag{4-7}$$

It remains to prove

$$\liminf_{N\to\infty}\mathrm{Tr}(U_R\gamma^{(2)}_{\widetilde{\Psi}_N})\geqslant 4\pi a\int\|u\|^4_{L^4}\,d\tilde{\mu}(u). \tag{4-8}$$

Note that (4-8) does not follow from (4-5) and Fatou's lemma easily because U_R depends on $R=R(N)$. We need to replace U_R by an operator bounded independently of N. Since $\tilde{h}$ has compact resolvent, for every $\Lambda\geqslant 1$, the projection

$$P_\Lambda:=\mathbb{1}(\tilde{h}\leqslant\Lambda)$$

has finite rank. Let us denote

$$\Pi:=\mathbb{1}_{\mathfrak{H}^2}-P_\Lambda^{\otimes 2}.$$

Since $U_R \geqslant 0$, we can apply the Cauchy–Schwarz inequality (3-9) with

$$X = P_\Lambda^{\otimes 2} U_R^{1/2} \quad \text{and} \quad Y = U_R^{1/2}\Pi$$

to obtain

$$\begin{aligned} U_R &= (P_\Lambda^{\otimes 2} + \Pi) U_R (P_\Lambda^{\otimes 2} + \Pi) \\ &= P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} + \Pi U_R \Pi + P_\Lambda^{\otimes 2} U_R \Pi + \Pi U_R P_\Lambda^{\otimes 2} \\ &\geqslant P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} - \delta^{-1} \Pi U_R \Pi - \delta P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} \end{aligned}$$

for all $\delta > 0$. Using the operator bound (3-4) and the fact that the $\frac{4}{5}$-th power is operator monotone [Bhatia 1997], we have

$$U_R(x_1 - x_2) \leqslant C \|U_R\|_{L^1} (1 - \Delta_1)^{4/5} (1 - \Delta_2)^{4/5} \leqslant C_{\varepsilon,s} (\tilde{h}_1)^{4/5} (\tilde{h}_2)^{4/5}. \tag{4-9}$$

Therefore,

$$P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} \leqslant C_{\varepsilon,s} \tilde{h}_1 \tilde{h}_2 \quad \text{and} \quad \Pi U_R \Pi \leqslant C_{\varepsilon,s} \Lambda^{-1/5} \tilde{h}_1 \tilde{h}_2.$$

Here in the second estimate, we have used $\mathbb{1}_{\mathfrak{H}} - P_\Lambda \leqslant \Lambda^{-1/5} (\tilde{h})^{1/5}$, which is a consequence of the definition of P_Λ. Thus

$$U_R - P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} \geqslant -C_{\varepsilon,s} (\delta^{-1} + \delta \Lambda^{-1/5}) \tilde{h}_1 \tilde{h}_2.$$

If we choose $\delta = \Lambda^{-1/10}$ and take the trace against $\gamma^{(2)}_{\tilde{\Psi}_N}$, then by the a priori estimate (4-3), we find

$$\operatorname{Tr}(U_R \gamma^{(2)}_{\tilde{\Psi}_N}) - \operatorname{Tr}\big(P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} \gamma^{(2)}_{\tilde{\Psi}_N}\big) \geqslant -C_{\varepsilon,s} \Lambda^{-1/10}. \tag{4-10}$$

On the other hand, from (4-9) and the definition of P_Λ, it follows that the operator norm $\|P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2}\|$ is bounded uniformly in N for fixed Λ. Therefore, the strong convergence (4-5) implies that

$$\lim_{N\to\infty} \left(\operatorname{Tr}\big(P_\Lambda^{\otimes 2} U_R P_\Lambda^{\otimes 2} \gamma^{(2)}_{\tilde{\Psi}_N}\big) - \int \big\langle (P_\Lambda u)^{\otimes 2}, U_R (P_\Lambda u)^{\otimes 2} \big\rangle \, d\tilde{\mu}(u) \right) = 0. \tag{4-11}$$

Since the left side of (4-7) is finite, every function u in the support of $d\tilde{\mu}$ belongs to the quadratic form domain $Q(\tilde{h})$ of $\tilde{h}$, and hence $P_\Lambda u \to u$ strongly in $Q(\tilde{h})$. Using the continuous embeddings $Q(\tilde{h}) \subset H^1 \subset L^4$, we get

$$\lim_{\Lambda\to\infty} \lim_{R\to 0} \big\langle (P_\Lambda u)^{\otimes 2}, U_R (P_\Lambda u)^{\otimes 2} \big\rangle = \lim_{\Lambda\to\infty} \|P_\Lambda u\|_{L^4}^4 = \|u\|_{L^4}^4.$$

By Fatou's lemma,

$$\liminf_{\Lambda\to\infty} \liminf_{N\to\infty} \int \big\langle (P_\Lambda u)^{\otimes 2}, U_R (P_\Lambda u)^{\otimes 2} \big\rangle \, d\tilde{\mu}(u) \geqslant 4\pi a \int \|u\|_{L^4}^4 \, d\tilde{\mu}(u). \tag{4-12}$$

The desired convergence (4-8) follows from (4-10), (4-11) and (4-12). □

Remark 4.2 (Mean-field approximation with two-body interactions).
From the preceding proposition, we obtain easily the convergence (2-11) mentioned in Section 2 because $\widetilde{H}_N \leqslant K_N$. In fact, K_N satisfies the second moment estimate (3-11) (see Remark 3.3), and hence (2-11) can be proved directly. In particular, the method can be used to derive the energy asymptotics when the interaction potential is given by (1-5); for $\beta < \frac{2}{3}$, Step 1 (and thus also Step 3) in the proof are not needed.

One can also obtain some explicit error estimate in Proposition 4.1 and (2-11) by using a quantitative version of the quantum de Finetti theorem as in [Lewin et al. 2015a, Lemma 3.4].

4B. ***Convergence of ground-state energy.*** We now conclude the proof of the convergence of the ground-state energy. There are two things left to do: remove the high momentum cut-off in the final effective functional, and relax the additional assumption (2-6).

Proof of energy convergence (1-9). The upper bound in (1-9) was proved in [Seiringer 2003]. The proof of the lower bound is divided into three steps.

Step 1. We start with the simple case when the extra condition (2-6) holds true. Recall that we are choosing

$$N^{-1/2} \gg R = R(N) \gg N^{-2/3}.$$

From Lemma 2.1 and Proposition 4.1, it follows that for every $1 > \varepsilon > 0$ and $s > 0$,

$$\liminf_{N\to\infty} \frac{\inf\sigma(H_N)}{N} \geqslant \liminf_{N\to\infty}\left(\frac{\inf\sigma(\widetilde{H}_N)}{N} + \kappa_{\varepsilon,s}\right) = e_{\mathrm{NL}}(\varepsilon,s) + \kappa_{\varepsilon,s}.$$

Thus to obtain the lower bound in (1-9), it remains to show that

$$\lim_{\varepsilon\to 0}\lim_{s\to\infty}(e_{\mathrm{NL}}(\varepsilon,s)+\kappa_{\varepsilon,s}) = e_{\mathrm{GP}}. \tag{4-13}$$

The upper bound in (4-13) is trivial as $\mathcal{E}_{\mathrm{NL}}(u)+\kappa_{\varepsilon,s} \leqslant \mathcal{E}_{\mathrm{GP}}(u)$. The lower bound in (4-13) can be done by a standard compactness argument provided in [Lieb and Seiringer 2006]. We recall this here for the reader's convenience. Let $u_{\varepsilon,s}$ be a ground state for $e_{\mathrm{NL}}(\varepsilon,s)$, namely

$$e_{\mathrm{NL}}(\varepsilon,s) = \langle u_{\varepsilon,s}, \tilde{h}u_{\varepsilon,s}\rangle + (1-\varepsilon)^2 4\pi a \int |u_{\varepsilon,s}|^4.$$

From (3-1), it follows that $\langle u_{\varepsilon,s}, (-\Delta+V)u_{\varepsilon,s}\rangle$ is bounded uniformly in s. Since $-\Delta+V$ has compact resolvent, for every given $\varepsilon>0$, there exists a subsequence $s_j\to\infty$ such that u_{ε,s_j} converges strongly in L^2 and pointwise (in both p-space and x-space) to a function u_ε. By Fatou's lemma, we have

$$\liminf_{j\to\infty}\int |u_{\varepsilon,s_j}(x)|^4\,dx \geqslant \int |u_\varepsilon(x)|^4\,dx,$$

$$\liminf_{j\to\infty}\int p^2(1-\theta_{s_j}(p))|\hat{u}_{\varepsilon,s_j}(p)|^2\,dp \geqslant \int p^2|\hat{u}_\varepsilon(p)|^2\,dp.$$

Next, using (2-6) as before, we have

$$\varepsilon p^2 + pA + Ap + |A|^2 + V + C_\varepsilon \geqslant 0$$

for some $C_\varepsilon \geqslant 0$. Using Fatou's lemma again and the strong convergence in L^2, we deduce

$$\liminf_{j\to\infty}\big\langle u_{\varepsilon,s_j}, \big(\varepsilon p^2 + pA + Ap + |A|^2 + V + \kappa_{\varepsilon,s}\big)u_{\varepsilon,s_j}\big\rangle \geqslant \big\langle u_\varepsilon, \big(\varepsilon p^2 + pA + Ap + |A|^2 + V\big)u_\varepsilon\big\rangle.$$

Combining these estimates, we get

$$\liminf_{j\to\infty}(e_{\mathrm{NL}}(\varepsilon,s_j)+\kappa_{\varepsilon,s_j}) \geqslant \langle u_\varepsilon, hu_\varepsilon\rangle + (1-\varepsilon)^2 4\pi a\int |u_\varepsilon|^4 \geqslant (1-\varepsilon)^2 e_{\mathrm{GP}}.$$

Taking $\varepsilon\to 0$, we obtain the lower bound in (4-13).

Step 2. From now on we do not assume (2-6). Let us introduce the Hamiltonian

$$H_{M,N} := \sum_{j=1}^{M} h_j + \sum_{1\leqslant i<j\leqslant M} w_N(x_i - x_j)$$

and denote by $E(M,N)$ its (bosonic) ground-state energy. In this step, we will prove the lower bound in (1-9) using the additional assumption

$$E(N,N) - E(N-1,N) \leqslant C. \tag{4-14}$$

We will find a function $f : \mathbb{R}^3 \to \mathbb{R}_+$ growing faster than $|A|$, namely

$$\lim_{|x|\to\infty} \frac{|A(x)|}{f(x)} = 0, \tag{4-15}$$

such that for a ground state Ψ_N for H_N, we have

$$\langle \Psi_N, f^2(x_1)\Psi_N\rangle \leqslant C. \tag{4-16}$$

Once this is achieved, we get

$$\frac{\inf\sigma(H_N)}{N} \geqslant \frac{\inf\sigma\big(H_N + \eta\sum_{j=1}^N f^2(x_j)\big)}{N} - C\eta$$

for every $\eta > 0$. Since the growth condition (2-6) holds true with V replaced by $V + \eta f^2$, we can apply the result in Step 1 to the Hamiltonian

$$H_N + \eta \sum_{j=1}^{N} f^2(x_j)$$

for every given $\eta > 0$. Then the lower bound in (1-9) follows by taking $\eta \to 0$.

Now we find such a function f. We will establish a simple binding inequality using an idea in [Lieb 1984]. From the ground-state equation $H_{N,N}\Psi_N = E(N,N)\Psi_N$, it follows that

$$E(N,N)\langle \Psi_N, f^2(x_N)\Psi_N\rangle = \Re\langle \Psi_N, f^2(x_N)H_{N,N}\Psi_N\rangle. \tag{4-17}$$

By the variational principle and (4-14), we have

$$H_{N,N} - h_N \geqslant H_{N-1,N} \geqslant E(N-1,N) \geqslant E(N,N) - C.$$

Note that $f^2(x_N)$ commutes with all terms in the latter inequality. If f is bounded and sufficiently regular, we have the IMS-type formula

$$\tfrac{1}{2}(f^2h + hf^2) = fhf - |\nabla f|^2 \geqslant Vf^2 - |\nabla f|^2, \tag{4-18}$$

and we deduce from (4-17) that

$$\big\langle \Psi_N, \big(V(x_N)f^2(x_N) - |\nabla f(x_N)|^2 - Cf^2(x_N)\big)\Psi_N\big\rangle \leqslant 0. \tag{4-19}$$

Note that if we choose $f(x) = e^{b|x|}$ for some constant $b > 0$, then (4-15) follows from the assumption (1-2). Moreover, heuristically, (4-16) follows from (4-19) as Vf^2 grows faster than $|\nabla f|^2 + Cf^2$. To make this idea rigorous, let us apply (4-19) with $f(x)$ replaced by

$$g_r(x) = \exp\bigl[b\bigl[r - \bigl||x| - r\bigr|\bigr]_+\bigr].$$

Note that $g_r(x) = e^{b|x|}$ when $|x| \leqslant r$ and $g_r(x) = 1$ when $|x| \geqslant 2r$. We can thus apply (4-18) to g_r.

Moreover,

$$\begin{aligned} Vg_r^2 - |\nabla g_r|^2 - Cg_r^2 &\geqslant (V - b^2 - C)g_r^2 \\ &\geqslant g_r^2 - (b^2 + C + 1)g_r^2 \mathbb{1}(V \leqslant b^2 + C + 1) \\ &\geqslant g_r^2 - C_0 \end{aligned}$$

for some constant C_0 independent of $r > 0$. Here we have used the fact that $g_r^2 \mathbb{1}(V \leqslant b^2 + C + 1)$ is bounded independently of $r > 0$, which follows from the assumption $\lim_{|x|\to\infty} V(x) = +\infty$. Thus (4-19) gives us

$$\langle \Psi_N, g_r(x_N)\Psi_N\rangle \leqslant C_0$$

for all $r > 0$. Taking $r \to \infty$, we obtain (4-16) with $f(x) = e^{b|x|}$.

Step 3. Now we explain how to remove the additional assumption (4-14). This can be done by following the strategy in [Lieb and Seiringer 2006], which we recall quickly below for the reader's convenience.

By choosing trial states $u^{\otimes N}$, we get the upper bound

$$E(N, N) \leqslant C_0 N$$

for some constant $C_0 > 2e_{\mathrm{GP}}$. For every $N \in \mathbb{N}$, we denote by $M = M(N)$ the largest integer $\leqslant N$ such that

$$E(M(N), N) - E(M(N) - 1, N) \leqslant C_0. \tag{4-20}$$

Then by the choice of $M(N)$, we obtain

$$E(N, N) - E(M(N), N) \geqslant (N - M(N))C_0. \tag{4-21}$$

We can find a subsequence $N_j \to \infty$ such that $M(N_j)/N_j \to \lambda \in [0, 1]$. Since (4-20) holds with $M = M(N_j)$, we can apply the result in Step 2 with w replaced by λw and find that

$$\liminf_{j\to\infty} \frac{E(M(N_j), N_j)}{M_j} \geqslant e_{\mathrm{GP}}(\lambda a) \geqslant \lambda e_{\mathrm{GP}}(a). \tag{4-22}$$

Here $e_{\mathrm{GP}}(\lambda a)$ is the Gross–Pitaevskii energy with a replaced by λa and the last inequality in (4-22) is obtained by simply ignoring part of the one-body energy in the corresponding Gross–Pitaevskii functional. From (4-21) and (4-22), it follows that

$$\begin{aligned} e_{\mathrm{GP}}(a) \geqslant \liminf_{j\to\infty} \frac{E(N_j, N_j)}{N_j} &\geqslant \liminf_{j\to\infty} \left(\frac{E(M(N_j), N_j)}{N_j} + C_0 \frac{N_j - M(N_j)}{N_j} \right) \\ &\geqslant \lambda^2 e_{\mathrm{GP}}(a) + C_0(1 - \lambda). \end{aligned}$$

Since

$$e_{\mathrm{GP}}(a) \leqslant \lambda^2 e_{\mathrm{GP}}(a) + 2(1 - \lambda)e_{\mathrm{GP}}(a)$$

and $C_0 > 2e_{\rm GP}(a)$, we must have $\lambda = 1$. Thus $M(N)/N \to 1$ for the whole sequence and

$$\liminf_{N\to\infty} \frac{E(N,N)}{N} = \liminf_{j\to\infty} \frac{E(N_j,N_j)}{N_j} \geqslant e_{\rm GP}(a).$$

This completes the proof of the energy convergence (1-9). □

4C. *Convergence of density matrices.* Now we prove the convergence of ground states in (1-10) by means of the Feynman–Hellmann principle. For $v \in L^2(\mathbb{R}^3)$ and $\ell \in \mathbb{N}$, we will perturb H_N by

$$S_{v,\ell} := \frac{\ell!}{N^{\ell-1}} \sum_{1\leqslant i_1<\cdots<i_\ell\leqslant N} |v^{\otimes \ell}\rangle\langle v^{\otimes\ell}|_{i_1,\dots,i_\ell}.$$

Here $|v^{\otimes \ell}\rangle\langle v^{\otimes\ell}|_{i_1,\dots,i_\ell}$ denotes the operator $|v^{\otimes \ell}\rangle\langle v^{\otimes\ell}|$ acting on the ℓ-body Hilbert space of the i_1-th,…, i_ℓ-th variables. We have the following extension of (1-9).

Lemma 4.3 (Energy lower bound for perturbed Hamiltonians).
We assume (1-2), (1-3) *and* (1-4). *For every* $v \in L^2(\mathbb{R}^3)$ *and* $\ell \in \mathbb{N}$, *we have*

$$\liminf_{N\to\infty} \frac{\inf\sigma(H_N - S_{v,\ell})}{N} \geqslant \inf_{\|u\|_{L^2}=1} (\mathcal{E}_{\rm GP}(u) - |\langle v,u\rangle|^{2\ell}). \tag{4-23}$$

Proof. We first work under the extra condition (2-6), and then explain how to remove it at the end. Let $1 > \varepsilon > 0$ and $s > 0$ and

$$N^{-1/2} \gg R = R(N) \gg N^{-2/3}.$$

Recall that from (2-2), we have

$$H_N - S_{v,\ell} \geqslant \widetilde{H}_N - S_{v,\ell} + N\kappa_{\varepsilon,s} - C_{\varepsilon,s}NR^2. \tag{4-24}$$

Let Φ_N be a ground state for $\widetilde{H}_N - S_{v,\ell}$. Since $\|S_{v,\ell}\|/N$ is bounded uniformly in N, (4-2) still holds true with $\widetilde{\Psi}_N$ replaced by Φ_N, namely

$$\langle \Phi_N, (\widetilde{H}_N)^k \Phi_N\rangle \leqslant (C_{\varepsilon,s}N)^k \tag{4-25}$$

for all $k \in \mathbb{N}$. Combining (4-25) with the three-body estimate in Lemma 3.4, we get the following analogue of (4-4):

$$\begin{aligned}\liminf_{N\to\infty} \frac{\inf\sigma(\widetilde{H}_N - S_{v,\ell})}{N} &= \liminf_{N\to\infty} \frac{\langle \Phi_N, (\widetilde{H}_N - S_{v,\ell})\Phi_N\rangle}{N} \\ &\geqslant \liminf_{N\to\infty} \big(\mathrm{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N}) + (1-\varepsilon)^2\,\mathrm{Tr}(U_R\gamma^{(2)}_{\widetilde{\Psi}_N}) - \mathrm{Tr}\big(|v^{\otimes\ell}\rangle\langle v^{\otimes\ell}|\gamma^{(\ell)}_{\Phi_N}\big)\big).\end{aligned} \tag{4-26}$$

Moreover, (4-25) and the second moment estimate (3-2) imply the a priori estimate $\langle \Phi_N, \tilde{h}_1\tilde{h}_2\Psi_N\rangle \leqslant C_{\varepsilon,s}$. Therefore, we can estimate the right side of (4-26) by proceeding exactly as in the proof of Proposition 4.1. More precisely, by the quantum de Finetti theorem, Theorem 2.2, we can find a Borel probability measure μ_Φ on the unit sphere $S\mathfrak{H}$ such that, up to a subsequence,

$$\lim_{N\to\infty} \mathrm{Tr}\left|\gamma^{(k)}_{\Phi_N} - \int |u^{\otimes k}\rangle\langle u^{\otimes k}|\, d\mu_\Phi(u)\right| = 0 \quad \forall k \in \mathbb{N}.$$

Using (4-6) with $\widetilde{\Psi}_N$ replaced by Φ_N and employing the fact that $|v^{\otimes \ell}\rangle\langle v^{\otimes \ell}|$ is bounded, we obtain

$$\liminf_{N\to\infty}\big(\operatorname{Tr}(\tilde{h}\gamma^{(1)}_{\widetilde{\Psi}_N}) + (1-\varepsilon)^2 \operatorname{Tr}(U_R\gamma^{(2)}_{\widetilde{\Psi}_N}) - \operatorname{Tr}\big(|v^{\otimes \ell}\rangle\langle v^{\otimes \ell}|\gamma^{(\ell)}_{\Phi_N}\big)\big) \\ \geqslant \int \left(\langle u, \tilde{h} u\rangle + (1-\varepsilon)^2 4\pi a \int |u|^4 - |\langle v, u\rangle|^{2\ell}\right) d\mu_\Phi(u). \quad (4\text{-}27)$$

From (4-24), (4-26) and (4-27), it follows that

$$\liminf_{N\to\infty} \frac{\inf\sigma(H_N - S_{v,\ell})}{N} \geqslant \inf_{\|u\|_{L^2}=1}\left(\langle u, \tilde{h} u\rangle + (1-\varepsilon)^2 4\pi a \int |u|^4 - |\langle v, u\rangle|^{2\ell}\right) + \kappa_{\varepsilon,s}.$$

The lower bound (4-23) follows by passing to the limits $s \to 0$ and then $\varepsilon \to 0$ as in the proof of (4-13).

To remove the assumption (2-6), we may use the argument in Section 4B. The only extra difficulty is that when dealing with the analogue of (4-17) with $H_{N,N}$ replaced by $H_{N,N} - S_{v,\ell}$, we have to take care of the operator $f^2|v\rangle\langle v| = |f^2 v\rangle\langle v|$, which may be unbounded as $f(x) = e^{b|x|}$ with $b > 0$ and v is merely in $L^2(\mathbb{R}^3)$. However, we can still proceed with all functions v in $L^2(\mathbb{R}^3)$ which have compact support. Then after obtaining the lower bound (4-23) with those nice functions v, we can extend the lower bound to all functions v in $L^2(\mathbb{R}^3)$ by a standard density argument. □

Now we are able to prove the convergence of density matrices.

Proof of state convergence (1-10). Let Ψ_N be an approximate ground state for H_N as in Theorem 1.1. For every $v \in L^2(\mathbb{R}^3)$ and $\ell \in \mathbb{N}$, from the upper bound in (1-9) and the lower bound in Lemma 4.3, we have

$$\begin{aligned}\limsup_{N\to\infty} \operatorname{Tr}\big(|v^{\otimes \ell}\rangle\langle v^{\otimes \ell}|\gamma^{(\ell)}_{\Psi_N}\big) &= \limsup_{N\to\infty}\left(\frac{\langle \Psi_N, H_N \Psi_N\rangle}{N} - \frac{\langle \Psi_N, (H_N - S_{v,\ell})\Psi_N\rangle}{N}\right) \\ &\leqslant \limsup_{N\to\infty}\left(\frac{\inf\sigma(H_N)}{N} - \frac{\inf\sigma(H_N - S_{v,\ell})}{N}\right) \\ &\leqslant e_{\mathrm{GP}} - \inf_{\|u\|_{L^2}=1}(\mathcal{E}_{\mathrm{GP}}(u) - |\langle v, u\rangle|^{2\ell}).\end{aligned}$$

Here v is not necessarily normalized. Therefore, we can replace v by $\lambda^{1/(2\ell)} v$ with $\lambda > 0$ and obtain

$$\limsup_{N\to\infty} \operatorname{Tr}\big(|v^{\otimes \ell}\rangle\langle v^{\otimes \ell}|\gamma^{(\ell)}_{\Psi_N}\big) \leqslant \frac{1}{\lambda}\big(e_{\mathrm{GP}} - \inf_{\|u\|_{L^2}=1}(\mathcal{E}_{\mathrm{GP}}(u) - \lambda|\langle v, u\rangle|^{2\ell})\big). \quad (4\text{-}28)$$

With given v and ℓ, for every $\lambda > 0$, let u_λ be a (normalized) minimizer for $u \mapsto \mathcal{E}_{\mathrm{GP}}(u) - \lambda|\langle v, u\rangle|^{2\ell}$. Since $\langle u_\lambda, h u_\lambda\rangle$ is bounded and h has compact resolvent, there exists a subsequence $\lambda_j \to 0$ such that u_{λ_j} converges to u_0 in L^2. By Fatou's lemma, u_0 is a minimizer of $\mathcal{E}_{\mathrm{GP}}(u)$. Moreover,

$$\limsup_{j\to\infty} \frac{1}{\lambda_j}\big(e_{\mathrm{GP}} - \inf_{\|u\|_{L^2}=1}(\mathcal{E}_{\mathrm{GP}}(u) - \lambda_j|\langle v, u\rangle|^{2\ell})\big) \\ \leqslant \limsup_{j\to\infty} \frac{1}{\lambda_j}\big(\mathcal{E}_{\mathrm{GP}}(u_{\lambda_j}) - (\mathcal{E}_{\mathrm{GP}}(u_{\lambda_j}) - \lambda_j|\langle v, u_{\lambda_j}\rangle|^{2\ell})\big) = |\langle v, u_0\rangle|^{2\ell}. \quad (4\text{-}29)$$

From (4-28) and (4-29), we conclude that for every $v \in L^2(\mathbb{R}^3)$ and $\ell \in \mathbb{N}$,

$$\limsup_{N\to\infty} \operatorname{Tr}\big(|v^{\otimes \ell}\rangle\langle v^{\otimes \ell}|\gamma^{(\ell)}_{\Psi_N}\big) \leqslant \sup_{u\in\mathcal{M}_{\mathrm{GP}}} |\langle v,u\rangle|^{2\ell}, \tag{4-30}$$

where $\mathcal{M}_{\mathrm{GP}}$ is the set of minimizers of $\mathcal{E}_{\mathrm{GP}}(u)$.

Note that also in [Lieb and Seiringer 2006], the upper bound (4-30) with $\ell = 1$ was proved, and from it, the convergence of the one-particle density matrices was deduced using an abstract argument of convex analysis. In the following, we will provide a simpler way to conclude the convergence of density matrices from (4-30), using the quantum de Finetti theorem. Indeed, by Theorem 2.2 as before, up to a subsequence of Ψ_N, there exists a Borel probability measure μ on the unit sphere $S\mathfrak{H}$ such that

$$\lim_{N\to\infty} \operatorname{Tr}\left|\gamma^{(k)}_{\Psi_N} - \int |u^{\otimes k}\rangle\langle u^{\otimes k}|\, d\mu(u)\right| = 0 \quad \forall k \in \mathbb{N}. \tag{4-31}$$

We will show that μ is supported on $\mathcal{M}_{\mathrm{GP}}$. From (4-30) and (4-31), we get

$$\int |\langle v,u\rangle|^{2k}\, d\mu(u) \leqslant \sup_{u\in\mathcal{M}_{\mathrm{GP}}} |\langle v,u\rangle|^{2k} \quad \forall v \in L^2(\mathbb{R}^3),\ k \in \mathbb{N}. \tag{4-32}$$

We assume for contradiction that there exists v_0 in the support of μ and $v_0 \notin \mathcal{M}_{\mathrm{GP}}$. We claim that we could then find $\delta \in (0, \frac{1}{2})$ such that

$$|\langle v,u\rangle| \leqslant 1 - 3\delta^2 \quad \forall u \in \mathcal{M}_{\mathrm{GP}},\ \forall v \in B, \tag{4-33}$$

where B is the set of all points in the support of μ within an L^2-distance less than δ from v_0. Indeed, if that were not the case, we would have two sequences strongly converging in L^2,

$$v_n \to v_0, \quad u_n \to u_0 \in \mathcal{M}_{\mathrm{GP}},$$

with $\|u_n - v_n\| \to 0$, and thus $v_0 \in \mathcal{M}_{\mathrm{GP}}$. Here we have used that $\mathcal{M}_{\mathrm{GP}}$ is a compact subset of $L^2(\mathbb{R}^3)$.

On the other hand, by the triangle inequality,

$$|\langle v,u\rangle| \geqslant \frac{\|u\|^2 + \|v\|^2 - \|u-v\|^2}{2} \geqslant 1 - 2\delta^2 \quad \forall u, v \in B. \tag{4-34}$$

Combining (4-32), (4-33) and (4-34), we find that

$$\begin{aligned}(\mu(B))^2(1-2\delta^2)^{2k} &\leqslant \int_B \int_B |\langle v,u\rangle|^{2k}\, d\mu(u)\, d\mu(v) \\ &\leqslant \int_B \sup_{u\in\mathcal{M}_{\mathrm{GP}}} |\langle v,u\rangle|^{2k}\, d\mu(v) \leqslant \mu(B)(1-3\delta^2)^{2k} \end{aligned}\tag{4-35}$$

for all $k \in \mathbb{N}$, and hence, taking $k \to \infty$, we have $\mu(B) = 0$. This, however, is a contradiction to the fact that v_0 belongs to the support of μ and μ is a Borel measure. Thus we conclude that μ is supported on $\mathcal{M}_{\mathrm{GP}}$ and the proof is complete. □

Acknowledgments

We thank Mathieu Lewin for helpful discussions. Phan Thành Nam is supported partially by the European Union's Seventh Framework Programme under REA grant agreement no. 291734. Nicholas Rougerie acknowledges financial support from the ANR (MaThoStaQ project, ANR-13-JS01-0005-01).

References

[Benedikter et al. 2015] N. Benedikter, G. de Oliveira, and B. Schlein, "Quantitative derivation of the Gross–Pitaevskii equation", *Comm. Pure Appl. Math.* **68**:8 (2015), 1399–1482. MR 3366749 Zbl 1320.35318

[Benedikter et al. 2016] N. Benedikter, M. Porta, and B. Schlein, *Effective evolution equations from quantum dynamics*, SpringerBriefs in Mathematical Physics **7**, Springer, Cham, 2016. MR 3382225 Zbl 06492716

[Bhatia 1997] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics **169**, Springer, New York, 1997. MR 98i:15003 Zbl 0863.15001

[Chen et al. 2015] T. Chen, C. Hainzl, N. Pavlović, and R. Seiringer, "Unconditional uniqueness for the cubic Gross–Pitaevskii hierarchy via quantum de Finetti", *Comm. Pure Appl. Math.* **68**:10 (2015), 1845–1884. MR 3385343 Zbl 1326.35332

[Chiribella 2011] G. Chiribella, "On quantum estimation, quantum cloning and finite quantum de Finetti theorems", pp. 9–25 in *Theory of quantum computation, communication, and cryptography*, edited by W. van Dam et al., Lecture Notes in Comput. Sci. **6519**, Springer, Berlin, 2011. MR 2012i:81019 Zbl 1310.81034

[Christandl et al. 2007] M. Christandl, R. König, G. Mitchison, and R. Renner, "One-and-a-half quantum de Finetti theorems", *Comm. Math. Phys.* **273**:2 (2007), 473–498. MR 2008m:81100 Zbl 1126.81032

[Dyson 1957] F. J. Dyson, "Ground-state energy of a hard-sphere gas", *Phys. Rev.* **106**:1 (1957), 20–26. Zbl 0077.23503

[Erdős and Yau 2001] L. Erdős and H.-T. Yau, "Derivation of the nonlinear Schrödinger equation from a many body Coulomb system", *Adv. Theor. Math. Phys.* **5**:6 (2001), 1169–1205. MR 2004c:82075 Zbl 1014.81063

[Erdős et al. 2007] L. Erdős, B. Schlein, and H.-T. Yau, "Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems", *Invent. Math.* **167**:3 (2007), 515–614. MR 2007m:81258 Zbl 1123.35066

[Erdős et al. 2009] L. Erdős, B. Schlein, and H.-T. Yau, "Rigorous derivation of the Gross–Pitaevskii equation with a large interaction potential", *J. Amer. Math. Soc.* **22**:4 (2009), 1099–1156. MR 2010g:82041 Zbl 1207.82031

[Erdős et al. 2010] L. Erdős, B. Schlein, and H.-T. Yau, "Derivation of the Gross–Pitaevskii equation for the dynamics of Bose-Einstein condensate", *Ann. of Math.* (2) **172**:1 (2010), 291–370. MR 2011g:82068 Zbl 1204.82028

[Hudson and Moody 1976] R. L. Hudson and G. R. Moody, "Locally normal symmetric states and an analogue of de Finetti's theorem", *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **33**:4 (1976), 343–351. MR 53 #1280 Zbl 0304.60001

[Lewin et al. 2014] M. Lewin, P. T. Nam, and N. Rougerie, "Derivation of Hartree's theory for generic mean-field Bose systems", *Adv. Math.* **254** (2014), 570–621. MR 3161107 Zbl 1316.81095

[Lewin et al. 2015a] M. Lewin, P. T. Nam, and N. Rougerie, "The mean-field approximation and the non-linear Schrödinger functional for trapped Bose gases", *Trans. Amer. Math. Soc.* (online publication October 2015).

[Lewin et al. 2015b] M. Lewin, P. T. Nam, and N. Rougerie, "Remarks on the quantum de Finetti theorem for bosonic systems", *Appl. Math. Res. Express. AMRX* 1 (2015), 48–63. MR 3335056 Zbl 1310.81169

[Lieb 1984] E. H. Lieb, "Bound on the maximum negative ionization of atoms and molecules", *Phys. Rev. A* **29**:6 (1984), 3018–3028.

[Lieb and Seiringer 2002] E. H. Lieb and R. Seiringer, "Proof of Bose–Einstein condensation for dilute trapped gases", *Phys. Rev. Lett.* **88**:17 (2002), paper ID 170409.

[Lieb and Seiringer 2006] E. H. Lieb and R. Seiringer, "Derivation of the Gross–Pitaevskii equation for rotating Bose gases", *Commun. Math. Phys.* **264**:2 (2006), 505–537. Zbl 1233.82004

[Lieb and Yngvason 1998] E. H. Lieb and J. Yngvason, "Ground state energy of the low density Bose gas", *Phys. Rev. Lett.* **80**:12 (1998), 2504–2507.

[Lieb et al. 2000] E. H. Lieb, R. Seiringer, and J. Yngvason, "Bosons in a trap: a rigorous derivation of the Gross–Pitaevskii energy functional", *Phys. Rev. A* **61**:4 (2000), paper ID 043602.

[Lieb et al. 2005a] E. H. Lieb, R. Seiringer, and J. P. Solovej, "Ground-state energy of the low-density Fermi gas", *Phys. Rev. A* **71**:5 (2005), paper ID 053605.

[Lieb et al. 2005b] E. H. Lieb, R. Seiringer, J. P. Solovej, and J. Yngvason, *The mathematics of the Bose gas and its condensation*, Oberwolfach Seminars **34**, Birkhäuser, Basel, 2005. MR 2006e:82001 Zbl 1104.82012

[Lieb et al. 2005c] E. H. Lieb, R. Seiringer, and J. Yngvason, "Justification of c-number substitutions in bosonic Hamiltonians", *Phys. Rev. Lett.* **94**:8 (2005), paper ID 080401.

[Pickl 2015] P. Pickl, "Derivation of the time dependent Gross–Pitaevskii equation with external fields", *Rev. Math. Phys.* **27**:1 (2015), paper ID 1550003. MR 3317556 Zbl 1320.35328

[Rougerie 2015] N. Rougerie, "De Finetti theorems, mean-field limits and Bose–Einstein condensation", lecture notes, 2015. arXiv 1506.05263

[Seiringer 2003] R. Seiringer, "Ground state asymptotics of a dilute, rotating gas", *J. Phys. A* **36**:37 (2003), 9755–9778. MR 2004h:82016 Zbl 1151.82401

[Størmer 1969] E. Størmer, "Symmetric states of infinite tensor products of C^*-algebras", *J. Functional Analysis* **3** (1969), 48–68. MR 39 #3327 Zbl 0167.43403

Received 30 Apr 2015. Revised 8 Sep 2015. Accepted 6 Jan 2016.

PHAN THÀNH NAM: pnam@ist.ac.at
Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

NICOLAS ROUGERIE: nicolas.rougerie@grenoble.cnrs.fr
Université Grenoble 1 and Centre National de la Recherche Scientifique, Laboratoire de Physique et Modélisation des Milieux Condensés (UMR 5493), BP 166, 38042 Grenoble, France

ROBERT SEIRINGER: robert.seiringer@ist.ac.at
Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.487

NONTRANSVERSAL INTERSECTION OF FREE AND FIXED BOUNDARIES FOR FULLY NONLINEAR ELLIPTIC OPERATORS IN TWO DIMENSIONS

EMANUEL INDREI AND ANDREAS MINNE

In the study of classical obstacle problems, it is well known that in many configurations, the free boundary intersects the fixed boundary tangentially. The arguments involved in producing results of this type rely on the linear structure of the operator. In this paper, we employ a different approach and prove tangential touch of free and fixed boundaries in two dimensions for fully nonlinear elliptic operators. Along the way, several n-dimensional results of independent interest are obtained, such as BMO-estimates, $C^{1,1}$-regularity up to the fixed boundary, and a description of the behavior of blow-up solutions.

1. Introduction

Optimal interior regularity results have recently been obtained for solutions to fully nonlinear free boundary problems [Figalli and Shahgholian 2014; Indrei and Minne 2015] via methods inspired by [Andersson et al. 2013]. Under further thickness assumptions, these results imply C^1-regularity of the free boundary. However, a description of the dynamics on how the free boundaries intersect the fixed boundary has remained an open problem for at least a decade in the fully nonlinear setting (although partial results have been obtained in [Matevosyan and Markowich 2004] under strong density and growth assumptions involving the solutions and a homogeneity assumption on the operator). On the other hand, extensive work has been carried out to investigate this question for the classical problem

$$\begin{cases} \Delta u = \chi_{u>0} & \text{in } B_1^+, \\ u \geq 0 & \text{in } B_1^+, \\ u = 0 & \text{on } \{x_n = 0\} \end{cases}$$

and its variations [Apushkinskaya and Uraltseva 1995; Shahgholian and Uraltseva 2003; Matevosyan 2005; Andersson et al. 2006; Andersson 2007]. The conclusions have varied as a function of the boundary data, but in the homogeneous case, it has been shown that the free boundary touches the fixed boundary tangentially. Dynamics of this type have also been the object of study in the classical dam problem [Caffarelli and Gilardi 1980; Alt and Gilardi 1982], which is a mathematical model describing the filtration of water through a porous medium split into wet and dry parts via a free boundary.

The methods utilized in establishing the above-mentioned results strongly rely on the linear structure of the operator, e.g., in arguments involving Green's functions and monotonicity formulas. In particular, the Alt–Caffarelli–Friedman and Weiss monotonicity formulas are frequently applied — tools only available

MSC2010: primary 35JXX, 35QXX; secondary 49SXX.
Keywords: obstacle problem, tangential touch, fully nonlinear equations, nontransverse intersection, free boundary problem.

in the setting of linear operators in divergence form; see [Petrosyan et al. 2012, Chapter 8]. Thus the tangential touch problem for fully nonlinear operators requires a different approach.

In this article, we prove nontransversal intersection of free and fixed boundaries in two dimensions for the problem

$$\begin{cases} F(D^2u) = \chi_\Omega & \text{a.e. in } B_1^+, \\ u = 0 & \text{on } B_1', \end{cases}$$

where $\Omega = \big(\{u \neq 0\} \cup \{\nabla u \neq 0\}\big) \cap \{x_2 > 0\} \subset \mathbb{R}^2_+$ and the free boundary is $\mathbb{R}^2_+ \cap \partial\Omega$. The starting point of our method is to first consider functions $u \in W^{2,n}(B_1^+)$ satisfying

$$\begin{cases} F(D^2u) = 1 & \text{a.e. in } B_1^+ \cap \Omega, \\ |D^2u| \le K & \text{a.e. in } B_1^+ \backslash \Omega, \\ u = 0 & \text{on } B_1', \end{cases} \tag{1}$$

where $\Omega \subset B_1^+$ is an (unknown) open set, $K > 0$, F is C^1 and satisfies standard structural assumptions (see Section 3).

Since by assumption D^2u is bounded in the complement of Ω, it follows that $F(D^2u)$ is bounded in B_1^+ and u is a strong L^n-solution to a fully nonlinear elliptic equation with bounded right-hand side [Caffarelli et al. 1996]. Under our structural assumptions on F, we have that L^n-solutions are also viscosity solutions, and it follows that $u \in W^{2,p}_{\text{loc}}(B_1^+)$ for all $p < \infty$ [Petrosyan et al. 2012]. If $u \ge 0$ and $\Omega = \{u \neq 0\}$, then since $D^2u = 0$ a.e. in the set $\{u = 0\}$, the Hessian condition in (1) is trivially satisfied; thus, (1) encodes the classical obstacle problem and likewise the equations $F(D^2u) = \chi_{u\neq 0}$, $F(D^2u) = \chi_{\nabla u \neq 0}$, and $F(D^2u) = \chi_{\{\nabla u\neq 0\}\cup\{u\neq 0\}}$ via the appropriate selection of Ω.

A heuristic description of our strategy is as follows: We consider

$$M := \limsup_{|x|\to 0} \frac{1}{x_n} \sup_{e \in \mathbb{S}^{n-2}\cap e_n^\perp} \partial_e u(x).$$

By extending interior $C^{1,1}$-results (see Section 3), it follows that M is finite, and we extract information on the nature of blow-up solutions by considering possible values for M. In particular, if $\{\nabla u \neq 0\} \cap \{x_n > 0\} \subset \Omega$ and the origin is a contact point, we show that either all blow-ups are of the form bx_n^2 if $M = 0$, or there is a sequence producing a blow-up having the form $ax_1x_n + bx_n^2$ if $M \neq 0$ (Theorem 2.1).

We then show that in $\mathbb{R}^2_+$, if $\Omega = \big(\{u \neq 0\} \cup \{\nabla u \neq 0\}\big) \cap \{x_2 > 0\}$ and $ax_1x_n + bx_n^2$ is a blow-up solution, then $\partial(\text{Int}\{u = 0\})$ stays away from the origin (Lemma 2.2) and this enables us to prove that blow-ups at the origin are unique (Theorem 2.4). Thereafter, a standard argument readily yields nontransversal intersection of the free and fixed boundaries at contact points (Theorem 2.5).

The rest of the paper is organized as follows: in the remainder of this section, we set up the problem and discuss relevant notation; Section 2 is the core of the paper where we rigorously develop the heuristics described above; Section 3 is devoted to the $C^{1,1}$-regularity up to the boundary of solutions, which follows as in [Indrei and Minne 2015] once a suitable BMO result is established. The results of Section 3 are used in Section 2. We have chosen to reverse the logical ordering of these sections in order to make the tangential touch section more accessible.

Setup and notation. We study fully nonlinear elliptic partial differential equations of the form

$$\begin{cases} F(D^2u, x) = f(x) & \text{a.e. in } B_1^+ \cap \Omega, \\ |D^2u| \le K & \text{a.e. in } B_1^+ \setminus \Omega, \\ u = 0 & \text{on } B_1', \end{cases} \tag{2}$$

where $u : B_1^+ \to \mathbb{R}$ is assumed to be in $W^{2,n}(B_1^+)$, $\Omega \subseteq B_1^+$ is an open set, $B_1(x) = \{x \in \mathbb{R}^n : |x| < 1\}$, $B_r^+(x) = B_r(x) \cap \{x_n > 0\}$, $B_r'(x) = B_r(x) \cap \{x_n = 0\}$, and $B_r = B_r(0)$.

Furthermore, F is assumed to satisfy the following structural conditions:

(H1) $F(0, x) \equiv 0$.

(H2) F is uniformly elliptic with ellipticity constants λ_0, $\lambda_1 > 0$ such that

$$\mathcal{P}^-(M - N) \le F(M, x) - F(N, x) \le \mathcal{P}^+(M - N) \quad \forall x \in B_1^+,$$

where M and N are symmetric matrices and $\mathcal{P}^\pm$ are the Pucci operators

$$\mathcal{P}^-(M) := \inf_{\lambda_0 \operatorname{Id} \le N \le \lambda_1 \operatorname{Id}} \operatorname{Tr} NM \quad \text{and} \quad \mathcal{P}^+(M) := \sup_{\lambda_0 \operatorname{Id} \le N \le \lambda_1 \operatorname{Id}} \operatorname{Tr} NM.$$

(H3) $F(\cdot, x)$ is concave or convex for all $x \in B_1^+$.

(H4)
$$|F(M, x) - F(M, y)| \le \bar{C}(|M| + 1)|x - y|^{\bar{\alpha}}$$

for some $\bar{\alpha} \in (0, 1]$ and x, $y \in B_1^+$.

Moreover, let

$$\beta(x, x^0) := \sup_{M \in \mathcal{S}} \frac{|F(M, x) - F(M, x^0)|}{|M| + 1},$$

where $\mathcal{S}$ is the space of $n \times n$ symmetric real valued matrices.

Points in $\mathbb{R}^n$ are generally denoted by x, x^0, y etc., while subscripts are used for components, i.e., $x = (x_1, \dots, x_n)$, scalar sequences, and functions. The notation x' is used for $(n-1)$-dimensional vectors. Similarly, ∇ and ∇' will be used, respectively, for the gradient and the gradient with respect to the first $n - 1$ variables. We will also use the following:

$\mathbb{R}^n_+$	is the upper half space $\{x \in \mathbb{R}^n : x_n > 0\}$;
Ω	is an open set in $\mathbb{R}^n_+$;
Γ	is the set $\mathbb{R}^n_+ \cap \partial\Omega$;
Γ_i	is the set $\mathbb{R}^n_+ \cap \partial \operatorname{Int}\{u = 0\}$;
$B_r(x^0)$	is the open ball $\{x \in \mathbb{R}^n : \lvert x - x^0\rvert < r\}$;
$B_r^+(x^0)$	is the truncated open ball $\{x \in \mathbb{R}^n : \lvert x - x^0\rvert < r,\ x_n > 0\}$;
$\partial B_r^+(x^0)$	is the topological boundary of $B_r^+(x^0)$ in $\mathbb{R}^n$;

B'_r is the ball $\{x' \in \mathbb{R}^{n-1} : |x'| < r\}$;

$\mathbb{S}^{n-1}$ is the $(n-1)$-sphere $\{x \in \mathbb{R}^n : |x| = 1\}$;

$e^\perp$ is the vector space orthogonal to $e \in \mathbb{S}^{n-1}$;

$C^{k,\alpha}(\Omega)$ denotes the usual Hölder space;

$C^{k,\alpha}_{\mathrm{loc}}(\Omega)$ denotes the local Hölder space;

$W^{k,p}(\Omega)$ denotes the usual Sobolev space.

The term "blow-ups of u at x^0" will be used for limits of the form

$$\lim_{j\to\infty} \frac{u(x^0 + r_j x)}{r_j^2},$$

where r_j is a sequence such that $r_j \to 0^+$ as $j \to \infty$, and $\mathrm{Int}\{u = 0\} = \{u = 0\}^\circ$ means the interior of the set $\{u = 0\} := \{x \in \mathbb{R}^n_+ : u(x) = 0\}$. Finally, $S(\psi)$ denotes the space of viscosity solutions corresponding to ψ and the ellipticity constants λ_0 and λ_1 in (H2); see [Caffarelli and Cabré 1995].

2. Main result

Our first result gives a natural dichotomy of blow-ups of solutions to (1) in any dimension.[1]

Theorem 2.1 (blow-up alternative). *Let u be a solution to* (1) *and suppose* $\{\nabla u \neq 0\} \cap \{x_n > 0\} \subset \Omega$, $0 \in \overline{\{u \neq 0\}}$, *and* $\nabla u(0) = 0$. *Then exactly one of the following holds*:

(i) *All blow-ups of u at the origin are of the form $u_0(x) = bx_n^2$ for some unique $b > 0$.*

(ii) *There exists a blow-up of u at the origin of the form*

$$u_0(x) = ax_1x_n + bx_n^2$$

for $a \neq 0$, $b \in \mathbb{R}$.

Proof. Firstly, since $u(x', 0) = 0$, it follows that $\partial_{x_i} u(x', 0) = 0$ for all $i \in \{1, \dots, n-1\}$. By $C^{1,1}$-regularity (Theorem 3.1), there is a constant $C > 0$ such that

$$\left| \frac{1}{x_n} \partial_{x_i} u(x', x_n) \right| = \left| \frac{1}{x_n} \big(\partial_{x_i} u(x', x_n) - \partial_{x_i} u(x', 0) \big) \right| \le C, \quad x_n > 0.$$

Define

$$M := \limsup_{\substack{|x|\to 0\\ x_n>0}} \frac{1}{x_n} \sup_{e\in\mathbb{S}^{n-2}\cap e_n^\perp} \partial_e u(x).$$

In particular, $0 \le M \le C < \infty$ and there exists a sequence $x^j \to 0$ with $x_n^j > 0$ and directions $e_{x^j} \in \mathbb{S}^{n-2}$ such that

$$\lim_{j\to\infty} \frac{1}{x_n^j} \partial_{e_{x^j}} u(x^j) = M.$$

[1] Regularity results from Section 3 will be utilized in the proof of Theorem 2.1.

Moreover, there exists $e \in \mathbb{S}^{n-2}$ such that (up to a subsequence) $e_{x^j} \to e$. Next note

$$\left|\frac{1}{x_n^j}\nabla' u(x^j)\cdot e - M\right| \leq \left|\frac{1}{x_n^j}\nabla' u(x^j)\cdot (e - e_{x^j})\right| + \left|\frac{1}{x_n^j}\nabla' u(x^j)\cdot e_{x^j} - M\right|$$
$$\leq C|e - e_{x^j}| + \left|\frac{1}{x_n^j}\nabla' u(x^j)\cdot e_{x^j} - M\right| \to 0$$

as $j \to \infty$. Thus, up to a rotation,

$$\lim_{j\to\infty} \frac{1}{x_n^j}\partial_{x_1} u(x^j) = M.$$

Thanks to uniform boundedness, consider a sequence $\{s_j\}$ such that $s_j \to 0^+$ and the corresponding blow-up procedure so that

$$u_j(x) := \frac{u(s_j x)}{s_j^2} \to u_0(x)$$

in $C^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^n_+)$ for any $\alpha \in [0, 1)$, and u_0 satisfies

$$\begin{cases} F(D^2 u_0) = 1 & \text{a.e. in } \mathbb{R}^n_+ \cap \Omega_0, \\ |\nabla u_0| = 0 & \text{in } \mathbb{R}^n_+ \setminus \Omega_0, \\ u = 0 & \text{on } \mathbb{R}^{n-1}_+, \end{cases} \tag{3}$$

where $\Omega_0 = \{\nabla u_0 \neq 0\} \cap \{x_n > 0\}$ (via nondegeneracy). The definition of M implies

$$M \geq \lim_j \left|\frac{\partial_{x_i} u(s_j x)}{s_j x_n}\right| = \lim_j \left|\frac{\partial_{x_i} u_j(x)}{x_n}\right| = \left|\frac{\partial_{x_i} u_0(x)}{x_n}\right| \tag{4}$$

for all $i \in \{1, \ldots, n-1\}$. In particular, let $v = \partial_{x_1} u_0$ so that in $\mathbb{R}^n_+$,

$$|v(x)| \leq M x_n. \tag{5}$$

If $M = 0$, then (4) implies $\partial_{x_i} u_0 = 0$ for all $i \in \{1, \ldots, n-1\}$ so that $u_0(x) = u_0(x_n)$. However, since $u_0(0) = |\nabla u_0(0)| = 0$, $0 \in \overline{\{u_0 \neq 0\}}$ and u_0 satisfies (3), the uniform ellipticity of F readily implies

$$u_0(x) = b x_n^2$$

for some unique $b > 0$. This shows that if $M = 0$, then any blow-up at the origin is of the form in (i).

Now suppose $M > 0$. In order to prove (ii), we cook up a specific blow-up: let $r_j := |x^j|$ (recall that $\{x^j\}$ is the sequence achieving the lim sup in the definition of M) so that as before

$$u_j(x) := \frac{u(r_j x)}{r_j^2} \to u_0(x)$$

in $C^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^n_+)$ for any $\alpha \in [0, 1)$, and u_0 satisfies (3), (4), and (5). Set $y^j = x^j / r_j \in \mathbb{S}^{n-1} \cap \{x_n > 0\}$, and note that along a subsequence, $y^j \to y \in \mathbb{S}^{n-1} \cap \{x_n \geq 0\}$. Moreover, by the choice of the sequence $\{x^j\}$

and the $C^{1,\alpha}$-convergence of u_j to u_0, if $y_n > 0$, then

$$\lim_j \frac{v(y^j)}{y_n^j} = \lim_j \frac{\partial_{x_1} u_j(y^j)}{y_n^j} = \lim_j \frac{\partial_{x_1} u(x^j)}{x_n^j} = M,$$

so that

$$v(y) = M y_n, \tag{6}$$

and note that (6) also holds if $y_n = 0$. We consider several possibilities, keeping in mind that $M > 0$.

Case 1: If $y \in \Omega_0$, then by differentiating (3), we get the elliptic equation

$$a_{ij}\partial_{ij}(v(x) - M x_n) = 0$$

for some measurable a_{ij}, and by (5), (6), and the maximum principle, it follows that $v(x) = M x_n$ in the connected component of Ω_0 containing y, say Ω_0^y. If there exists $x \in \partial\Omega_0^y \cap \{x_n > 0\}$, then $M x_n = v(x) = 0$, so we must have $M = 0$, a contradiction. Thus, $v(x) = M x_n$ in $\mathbb{R}^n_+$ and by integrating,

$$u_0(x) = M x_1 x_n + h(x_2, \dots, x_n).$$

Now, the $C^{2,\alpha}$-estimate up to the boundary given by the Krylov–Safonov theorem (see, e.g., Theorem 3.3) applied to $u_0(Rx)/R^2$ yields

$$\frac{|D^2 u_0(x) - D^2 u_0(y)|}{|x - y|^\alpha} \le \frac{C}{R^\alpha}, \quad x \ne y \in B_R^+,$$

and taking $R \to \infty$ implies that $D^2 u_0$ is a constant matrix, and thus h is a second-order polynomial. Since u_0 vanishes on $\{x_n = 0\}$, it follows that

$$h(x_2, \dots, x_n) = x_n \sum_{i \ne n} \alpha_i x_i + b x_n^2,$$

and so up to a rotation,

$$u_0(x) = a x_1 x_n + b x_n^2,$$

with a or $b \ne 0$.

Case 2: If $y \in \partial\Omega_0 \cap \{x_n > 0\}$, then $M y_n = v(y) = 0$, a contradiction.

Case 3: If $y \in \overline{\Omega}_0^c$, then for all but finitely many j, we have $y^j \in \Omega_0^c$ and since $\{\nabla u_0 \ne 0\} \subset \Omega_0$, it follows that $v(y^j) = 0$ if $j \ge N$ for some $N \in \mathbb{N}$. However, $y_n^j > 0$ and so

$$0 = \lim_j \frac{v(y^j)}{y_n^j} = M,$$

a contradiction.

Case 4: If $y \in \partial\Omega_0 \cap \{x_n = 0\}$, by differentiating (3) in Ω_0, it can be seen that for $r > 0$ (to be picked later), v satisfies

$$Lv = 0 \ \text{ in } B_r(y)^+ \cap \Omega_0,$$

where $L = F_{ij}(D^2u_0)\partial_{ij}$ is elliptic. Since $u_0 \in C^{1,1}(B_r^+(y))$, it follows that the $F_{ij}(D^2u_0)$ are bounded and measurable on $B_r^+(y)$.

We know that $Mx_n - v(x) \geq 0$ in $\mathbb{R}^n_+$, and if equality holds everywhere, $u_0(x) = ax_1x_n + bx_n^2$ just as in Case 1. If there is a point z where strict inequality holds, that is, $Mz_n - v(z) > 0$, then we can choose a ball $B_r^+(y)$ so that, by continuity of v, we have $v(x) < Mx_n$ in a neighborhood $B_s(z)$, where z is a boundary point of $B_r^+(y)$. Note that this strict inequality necessarily occurs on $\partial B_r^+(y) \cap \{x_n > 0\}$ since both v and Mx_n are zero on the hyperplane $\{x_n = 0\}$. Now choose a smooth nonnegative (but not identically zero) function ϕ supported on $B_s(z)$ small enough such that $Mx_n - \phi(x) \geq v(x)$ and $Mx_n - \phi(x) > 0$ in $\mathbb{R}^n_+$ (this can be done since $B_s(z)$ is some distance away from the hyperplane $\{x_n = 0\}$). Then if

$$\begin{cases} Lw = 0 & \text{in } B_r^+(y), \\ w = Mx_n - \phi & \text{on } \partial B_r^+(y), \end{cases}$$

we have that $w > 0$ in $B_r^+(y)$ by the strong maximum principle since $Mx_n - \phi(x) > 0$. In particular, $w > v = 0$ on $\partial\Omega$, and since $v \leq w$ on $\partial B_r^+(y)$, the strong maximum principle again gives $w > v$ in $B_r^+(y) \cap \Omega$. Note also by linearity that $w = Mx_n - h$, where h solves

$$\begin{cases} Lh = 0 & \text{in } B_r^+(y), \\ h = \phi & \text{on } \partial B_r^+(y). \end{cases}$$

Once more, the strong maximum principle shows that $h > 0$ in $B_r^+(y)$, so the boundary Harnack comparison principle implies that $cx_n \leq h(x)$ in $B_{r/2}^+(y)$, where $c > 0$ depends on ellipticity constants and ϕ. With this,

$$M = \lim_{j\to\infty} \frac{v(y^j)}{y_n^j} \leq \limsup_{\substack{x_n\to 0^+ \\ x\in B_{r/4}^+(y)}} \frac{w(x)}{x_n} \leq \lim_{\substack{x_n\to 0^+ \\ x\in B_{r/4}^+(y)}} \frac{Mx_n - cx_n}{x_n} = M - c,$$

a contradiction. □

The next lemma shows that in two dimensions, if (ii) in Theorem 2.1 occurs, then $\Gamma_i = \mathbb{R}^n_+ \cap \partial\,\mathrm{Int}\{u = 0\}$ stays away from the origin.

Lemma 2.2. *Let u be a solution to* (1) *with $\Omega = \big(\{u \neq 0\} \cup \{\nabla u \neq 0\}\big) \cap \{x_2 > 0\} \subset \mathbb{R}^2_+$. If there exists $\{r_j\} \subset \mathbb{R}^+$ such that $r_j \to 0$ as $j \to \infty$ and*

$$u_j(x) := \frac{u(r_jx)}{r_j^2} \to u_0(x) = ax_1x_2 + bx_2^2$$

in $C^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^n_+)$ as $j \to \infty$ for $a \neq 0$, $b \in \mathbb{R}$, then there exists $\delta \in (0,1)$ such that $B_\delta^+ \cap \Gamma_i = \varnothing$.

Proof. We may assume $a > 0$. Set $v_j := \partial_1 u_j$ and let $R > 2$, $\mu \in \big(0, \frac{1}{4}\big)$, and $\delta \in \big(0, \frac{1}{4}\big)$. Then select $j_0 = j_0(R, \mu, \delta) > 0$ such that for all $j \geq j_0$,

$$|\nabla u_j(x)| > 0, \quad x \in B_R^+ \setminus B_\delta^+, \tag{7}$$

$$v_j(x) > 0, \quad x \in B_R^+ \cap \{x_2 \geq \mu\} \tag{8}$$

(the two-dimensional setting is crucial for (7)). Consider $z \in \partial B_1 \cap \{x_2 = 0\}$ and note that

$$B^+_{3/4}(z) \subset B^+_R \setminus B^+_\delta.$$

Thanks to (7), u_j satisfies $F(D^2 u_j) = 1$ in $B^+_{3/4}(z)$ for all $j \geq j_0$. $C^{2,\alpha}$-estimates up to the boundary (see Theorem 3.3) imply

$$\sup_j \|u_j\|_{C^{2,\alpha}\left(B^+_{3/4}(z)\right)} < \infty.$$

Thus, along a subsequence, $v_j \to a x_2 =: v$ in $C^{0,1}$ ($C^{2,\alpha}$ is compactly contained in $C^{1,1}$) and so

$$c_j := \sup_{\substack{x,y \in B^+_{3/4}(z)\\ x \neq y}} \frac{|(v_j(x) - v_j(y)) - (v(x) - v(y))|}{|x-y|} \to 0.$$

In particular, since $v_j(x_1, 0) = v(x_1, 0) = 0$, it follows that

$$\frac{|v_j(x) - a x_2|}{x_2} \leq c_j,$$

and so

$$v_j(x) \geq (a - c_j) x_2.$$

Now we select j large such that $v_j(x) \geq 0$ on ∂B_1. Note that $L v_j = 0$ in $B^+_1 \cap \Omega(u_j)$, where L is an elliptic second-order operator obtained by differentiating (1). Indeed, u_j satisfies

$$\begin{cases} F(D^2 u_j) = 1 & \text{a.e. in } B^+_{1/r_j} \cap \Omega(u_j),\\ u_j = 0 & \text{on } B'_{1/r_j}, \end{cases}$$

where $\Omega(u_j)$ is the dilated set Ω/r_j, and without loss we may assume $r_j < \frac{1}{2}$.

Since v_j vanishes on $\partial\Omega(u_j)$ and is nonnegative on ∂B^+_1, the maximum principle implies $v_j > 0$ in $B^+_1 \cap \Omega(u_j)$ (note that v_j is not identically zero by (8)). If $\Gamma_i(u_j) \cap B^+_\delta \neq \varnothing$, consider a ball N in the interior of $\{u_j = 0\} \cap B^+_\delta$. For $t \in \mathbb{R}$, let $N_t = N + t e_1$. Note that by taking t negative, we can move N_t to the left so that eventually $N_t \subset B^+_1 \setminus B^+_\delta$. Consider the strip $S = \bigcup_{t \in \mathbb{R}} N_t$. The next claim is that there exists a ball in the set $(S \cap B^+_1) \setminus B^+_\delta$ such that $u_j \neq 0$ in this ball: if not, then for each point $z \in (S \cap B^+_1) \setminus B^+_\delta$, there exists a sequence $\{z_k\} \subset \{u_j = 0\}$ such that $z_k \to z$; by continuity, $u_j(z) = 0$, so $u_j = 0$ in $(S \cap B^+_1) \setminus B^+_\delta$, and therefore the gradient also vanishes there, a contradiction to (7). Denote the ball by $E \subset \Omega(u_j)$ and note that $u_j < 0$ on E since for each $z \in E$, there exists $t_z > 0$ such that $z + e_1 t_z \in \{u_j = 0\}$ and $v_j > 0$ in $B^+_1 \cap \Omega(u_j)$. Thus, $E \subset \Omega(u_j) \cap \{u_j < 0\}$. Now move E to the right until the first time it touches $\{u_j = 0\}$, and let y be the contact point.

If $\nabla u_j(y) = 0$, we immediately obtain a contradiction via Hopf's lemma. Thus we may assume $\nabla u_j(y) \neq 0$, which implies $y \in \Omega(u_j)$, whence $v_j(y) > 0$ (recall that $v_j > 0$ in $\Omega(u_j)$). By continuity, $v_j > 0$ in $B_r(y)$ for some $r > 0$, so in particular $v_j(y + t e_1) > 0$ for all $t > 0$ small. Since $\{y + t e_1 : t \in (0, r)\} \subset \Omega(u_j)$, we know $t_* := \sup\{t > 0 : y + t e_1 \in \Omega(u_j)\}$ is positive. Note that $y + t e_1$

will eventually enter N as t gets larger. However,

$$u_j(y+t_*e_1)-u_j(y)=\int_0^{t_*} v_j(y+se_1)\,ds>0,$$

and this implies $0=u_j(y+t_*e_1)>u_j(y)=0$, a contradiction. Thus $\Gamma_i(u_j)\cap B_\delta^+=\varnothing$ and the result follows. □

Before proving uniqueness of blow-ups and tangential touch, we require one more lemma.

Lemma 2.3. *Let u be a solution to* (1) *with $\Omega=\big(\{u\neq 0\}\cup\{\nabla u\neq 0\}\big)\cap\{x_n>0\}$. If $s\in(0,1]$ and $(B_s^+\setminus\Omega)^\circ=\varnothing$, then $|B_s^+\setminus\Omega|=0$.*

Proof. Since $u\in W^{2,n}(B_1^+)$, it follows that $D^2u=0$ a.e. on $B_s^+\setminus\Omega$. Let $Z:=\{D^2u=0\}\cap(B_s^+\setminus\Omega)$ and note that $|Z|=|B_s^+\setminus\Omega|$. Thus if $Z\subset(B_s^+\setminus\Omega)^\circ$, then the result follows. Let $x^0\in Z$ and suppose $x^0\notin(B_s^+\setminus\Omega)^\circ$. Then consider a sequence of points $x^j\to x^0$ such that $u(x^j)\neq 0$ and let $r_j:=|x^0-x^j|$. Nondegeneracy (see, e.g., [Indrei and Minne 2015, Lemma 3.1]) implies that for j large,

$$\sup_{\partial B_{r_j}(x^0)}\frac{u}{r_j^2}\geq c>0,$$

or in other words,

$$\sup_{\partial B_1(0)}\frac{u(x^0+r_jx)}{r_j^2}\geq c>0.$$

Now for each j large enough, let $y^j\in\partial B_1(0)$ be the element achieving the supremum in the previous expression; note that since

$$u(x^0)=|\nabla u(x^0)|=|D^2u(x^0)|=0,$$

we have

$$u(x^0+r_jy^j)=o(r_j^2),$$

a contradiction. □

Theorem 2.1, Lemma 2.2, and Lemma 2.3 imply uniqueness of blow-ups in two dimensions.

Theorem 2.4 (uniqueness of blow-ups). *Let u be a solution to* (1) *with $\Omega=\big(\{u\neq 0\}\cup\{\nabla u\neq 0\}\big)\cap\{x_2>0\}\subset\mathbb{R}_+^2$. If $0\in\overline{\{u\neq 0\}}$ and $\nabla u(0)=0$, then all blow-up limits u_0 of u at the origin are of the form*

$$u_0(x)=ax_1x_2+bx_2^2,$$

where $a,b\in\mathbb{R}$ with at least one of them nonzero.

Proof. We divide the proof into two cases.

Case 1: $0\in\overline{\Gamma}_i$. Lemma 2.2 implies the nonexistence of a blow-up u_0 of u of the form

$$ax_1x_2+bx_2^2,$$

$a\neq 0$, $b\in\mathbb{R}$, from which it follows that (i) holds in Theorem 2.1. Note that b is uniquely determined by the equation.

Case 2: $0 \notin \overline{\Gamma}_i$. In this case, there exists $\delta > 0$ such that $\Gamma_i \cap B_\delta^+ = \varnothing$. Since $0 \in \overline{\{u \neq 0\}}$ (by assumption), it follows that $B_\delta^+ \not\subset \{u = 0\}^\circ$ and as $\Gamma_i \cap B_\delta^+ = \varnothing$, we may conclude that $\{u = 0\}^\circ \cap B_\delta^+ = \varnothing$. Thus the hypotheses of Lemma 2.3 are satisfied, and by applying the lemma, we obtain that $F(D^2u) = 1$ a.e. in B_δ^+. Therefore $u \in C^{2,\alpha}(B_{\delta/2}^+)$ and the blow-up limit u_0 is uniquely given by

$$\begin{aligned}\lim_{r\to 0}\frac{u(rx)}{r^2} &= \lim_{r\to 0}\frac{u(0)+\nabla u(0)\cdot rx + \langle rx, D^2u(0)rx\rangle + o(r^2)}{r^2}\\ &= \langle x, D^2u(0)x\rangle = ax_1x_2 + bx_2^2.\end{aligned}$$

The last equality follows from the boundary condition. Furthermore, u_0 solves the same equation as u, so

$$F(D^2u_0) = F(D^2u(0)) = 1$$

and thus a and b cannot both be zero due to (H1). □

If blow-ups are unique and of the form given above, it is rather standard to show that the free boundary touches the fixed boundary tangentially (see, e.g., [Petrosyan et al. 2012, Chapter 8]). The proof is included for completeness.

Theorem 2.5 (tangential touch). *Let u be a solution to* (1) *with* $\Omega = \big(\{u \neq 0\} \cup \{\nabla u \neq 0\}\big) \cap \{x_2 > 0\} \subset \mathbb{R}^2_+$. *Then there exists a constant $r_0 > 0$ and a modulus of continuity $\omega_u(r)$ such that*

$$\Gamma(u) \cap B_{r_0}^+ \subset \big\{x : x_2 \leq \omega_u(|x|)|x|\big\}$$

if $0 \in \overline{\Gamma(u)}$, *where* $\Gamma(u) := \partial\Omega \cap \mathbb{R}^2_+$.

Proof. By Theorem 2.4, the blow-up of u at the origin is not identically zero and is given by $u_0(x) = ax_1x_2 + bx_2^2$. In particular, $\Gamma(u_0) = \varnothing$. It suffices to show that for any $\epsilon > 0$, there exists $\rho_\epsilon = \rho_\epsilon(u) > 0$ such that

$$\Gamma(u) \cap B_{\rho_\epsilon}^+ \subset B_{\rho_\epsilon}^+ \setminus \mathcal{C}_\epsilon,$$

where $\mathcal{C}_\epsilon := \{x_2 > \epsilon|x_1|\}$. Suppose not. Then there exists a solution u to (1) satisfying the hypotheses of the theorem and $\epsilon > 0$ such that for all $k \in \mathbb{N}$, there exists

$$x^k \in \Gamma(u) \cap B_{1/k}^+ \cap \mathcal{C}_\epsilon.$$

Let $r_k := |x^k|$ and $y^k := x^k/r_k \in \partial B_1 \cap \mathcal{C}_\epsilon$. Note that along a subsequence

$$y^k \to y \in \partial B_1 \cap \mathcal{C}_\epsilon.$$

Define

$$u_k(x) := \frac{u(r_kx)}{r_k^2}$$

so that $u_k \to u_0$ in $C^{1,\alpha}_{\text{loc}}(\mathbb{R}^n_+)$ (along a subsequence). In particular, $y \in \Gamma(u_0)$ by nondegeneracy, which contradicts that $\Gamma(u_0) = \varnothing$. □

3. $C^{1,1}$-regularity up to the boundary

We now show BMO-estimates as well as $C^{1,1}$-regularity up to the fixed boundary of solutions to (2).

Theorem 3.1 ($C^{1,1}$-regularity). *Let $f \in C^{\alpha}(B_1^+)$ be a given function and $\Omega \subseteq B_1^+$ a domain such that $u : B_1^+ \to \mathbb{R}$ is a $W^{2,n}$-solution of* (2). *Assume F satisfies* (H1)–(H4). *Then there exists a constant C depending on $\|u\|_{W^{2,n}(B_1^+)}$, $\|f\|_{C^{\alpha}(B_1^+)}$, and universal constants such that*

$$|D^2u| \le C \quad \text{a.e. in } B_{1/2}^+.$$

There are three key tools needed to prove this theorem. The first two are $C^{2,\alpha}$- and $W^{2,p}$-estimates up to the boundary for the following classical fully nonlinear problem

$$\begin{cases} F(D^2u, x) = f(x) & \text{a.e. in } B_1^+, \\ u = 0 & \text{on } B_1', \end{cases} \tag{9}$$

and the last involves BMO-estimates. The $C^{2,\alpha}$- and $W^{2,p}$-estimates are well-known [Wang 1992; Safonov 1994; Winter 2009; Krylov 1982]. We have been unable to find a reference for the BMO-estimates and thus provide a proof, which is an adaptation of the interior case. For convenience, we record the following estimates; see, e.g., [Winter 2009, Theorem 4.3; Safonov 1994, Theorem 7.1]. Recall the definition of β,

$$\beta(x, x^0) := \sup_{M \in \mathcal{S}} \frac{|F(M, x) - F(M, x^0)|}{|M| + 1}.$$

Theorem 3.2 ($W^{2,p}$-regularity). *Let u be a $W^{2,p}$-viscosity solution to* (9) *and $f \in L^p(B_1^+)$ for $n \le p \le \infty$. If $\beta(x^0, y) \le \beta_0$ in $B_r^+(x^0) \cap B_1^+$ for all $x^0 \in B_1^+$ and $0 < r \le r_0$, where β_0 and r_0 are universal constants, then $u \in W^{2,p}(B_{1/2}^+)$ and*

$$\|u\|_{W^{2,p}(B_{1/2}^+)} \le C\big(\|u\|_{L^\infty(B_1^+)} + \|f\|_{L^p(B_1^+)}\big),$$

where $C = C(n, \lambda_0, \lambda_1, \bar{\alpha}, \bar{C}, p) > 0$.

Theorem 3.3 ($C^{2,\alpha}$-regularity). *Let u be a $W^{2,n}$-viscosity solution to* (9) *and $f \in C^{\bar{\alpha}}(B_1^+)$. Then if $\beta(x^0, y) \le \beta_0$ in $B_r^+(x^0) \cap B_1^+$ for all $x^0 \in B_1^+$ and $0 < r \le r_0$, where β_0 and r_0 are universal constants, then $u \in C^{2,\alpha}(B_{1/2}^+)$ and*

$$\|u\|_{C^{2,\alpha}(B_{1/2}^+)} \le C\big(\|u\|_{L^\infty(B_1^+)} + \|f\|_{C^{\bar{\alpha}}(B_1^+)}\big),$$

where $C = C(n, \lambda_0, \lambda_1, \bar{\alpha}, \bar{C}) > 0$.

The next results are technical tools utilized in the proof of the BMO-estimate (i.e., Proposition 3.6). The first is an approximation lemma; see, e.g., [Wang 1992, Lemma 1.4].

Lemma 3.4 (approximation). *Let $\epsilon > 0$, $u \in W^{2,p}(B_1^+(x^0))$, and let v solve*

$$\begin{cases} F(D^2v, x^0) = a & \text{in } B_{1/2}^+(x^0), \\ v = u & \text{on } \partial B_{1/2}^+(x^0). \end{cases}$$

Then there exists $\delta > 0$ *and* $\eta > 0$ *such that if*

$$\beta(x, x^0) := \sup_{M \in \mathcal{S}} \frac{|F(M, x) - F(M, x^0)|}{|M| + 1} \le \eta$$

and $|f(x) - a| \le \delta$ *a.e. for* $f(x) := F(D^2u(x), x)$ *in* $B_1^+(x^0)$*, then*

$$|u - v| \le \epsilon \quad \text{in } B_{1/2}^+.$$

Lemma 3.5. *Let* $u \in W^{2,n}(B_1^+)$ *satisfy* $|F(D^2u(x), x)| \le \delta$ *a.e. in* B_1^+ *for* δ *as in Lemma 3.4. Moreover, assume* $|u| \le 1$ *and that* $\beta(x, y)$ *satisfies (H4). Then there exists a universal constant* $\rho > 0$ *and second-order polynomials* P_{k,x^0} *for any* $k \in \mathbb{N}_0$ *and* $x^0 \in B_{1/2}^+$ *such that*

$$\begin{aligned} |D^2 P_{k,x^0} - D^2 P_{k-1,x^0}| &\le C_0(n, \lambda_0, \lambda_1), \\ F(D^2 P_{k,x^0}, x^0) &= 0, \\ |u(x) - P_{k,x^0}(x)| &\le \rho^{2k} \quad \text{inside } B^+_{\min(\rho^k,1)}(x^0). \end{aligned}$$

Proof. For $k = 0$ and $k = -1$, the statement is true for $P_{k,x^0}(x) \equiv 0$ by assumption (recall (H1)). If we assume it is true up to some k, define

$$\begin{aligned} u_k &:= \frac{u(\rho^k x + x^0) - P_{k,x^0}(\rho^k x + x^0)}{\rho^{2k}}, \\ F_k(M, x) &:= F\big(M + D^2 P_{k,x^0}, \rho^k x + x^0\big), \quad x \in B_1 \cap \{x_n > -x_n^0/\rho^k\}. \end{aligned}$$

Then $|F_k(D^2u_k, x)| = \big|F\big((D^2u)(\rho^k x + x^0), \rho^k x + x^0\big)\big| \le \delta$ a.e. Also,

$$\begin{aligned} \beta_k(x, 0) &= \sup_{M \in \mathcal{S}} \frac{|F_k(M, x) - F_k(M, 0)|}{|M| + 1} \\ &= \sup_{M \in \mathcal{S}} \frac{\big|F\big(M + D^2 P_{k,x^0}, \rho^k x + x^0\big) - F\big(M + D^2 P_{k,x^0}, x^0\big)\big|}{|M| + 1} \\ &= \sup_{M \in \mathcal{S}} \frac{|F(M, \rho^k x + x^0) - F(M, x^0)|}{|M - D^2 P_{k,x^0}| + 1} \\ &= \sup_{M \in \mathcal{S}} \frac{|F(M, \rho^k x + x^0) - F(M, x^0)|}{|M| + 1} \frac{|M| + 1}{|M - D^2 P_{k,x^0}| + 1} \\ &\le \beta(\rho^k x + x^0, x^0) \sup_{M \in \mathcal{S}} \frac{|M| + 1}{|M - D^2 P_{k,x^0}| + 1} \\ &\le \bar{C} \rho^{\bar{\alpha} k} \sup_{M \in \mathcal{S}} \frac{|M| + 1}{\big||M| - |D^2 P_{k,x^0}|\big| + 1} \\ &\le \bar{C} \rho^{\bar{\alpha} k} \big(|D^2 P_{k,x^0}| + 1\big), \end{aligned}$$

where the last inequality follows from a calculation of the maximum of the function

$$\frac{x+1}{|x-a|+1}, \quad x, a \geq 0.$$

However, from the induction hypothesis,

$$|D^2 P_{k,x^0}| \leq \sum_{j=1}^{k} \big|D^2 P_{j-1,x^0} - D^2 P_{j,x^0}\big| \leq C_0 k,$$

so

$$\bar{C}\rho^{\bar{\alpha}k}\big(|D^2 P_{k,x^0}| + 1\big) \leq \bar{C}\rho^{\bar{\alpha}k} C_0 k \leq \eta$$

if ρ is chosen small enough (depending only on universal constants) and η as in Lemma 3.4. Thus $|v_k - u_k| \leq \epsilon$ in $B_{1/2} \cap \{x : x_n > -x_n^0/\rho^k\}$ by Lemma 3.4, where v_k solves

$$\begin{cases} F_k(D^2 v_k, 0) = 0 & \text{in } B_{1/2} \cap \{x : x_n > -x_n^0/\rho^k\}, \\ v_k = u_k & \text{on } \partial\big(B_{1/2} \cap \{x : x_n > -x_n^0/\rho^k\}\big). \end{cases}$$

Since

$$\|v_k\|_{L^\infty(B_{1/2}\cap\{x:x_n>-x_n^0/\rho^k\})} \leq \|u_k\|_{L^\infty(B_{1/2}\cap\{x:x_n>-x_n^0/\rho^k\})} \leq 1$$

by the maximum principle, Theorem 3.3 gives

$$\|v_k\|_{C^{2,\alpha}(B_{1/4}\cap\{x:x_n>-x_n^0/\rho^k\})} \leq C_0. \tag{10}$$

Now define $\hat{P}_{k,x^0}$ as the second-order Taylor expansion of v_k at the origin, and note that $F_k(D^2\hat{P}_{k,x^0}, 0) = F_k(D^2 v_k(0), 0) = 0$. Then

$$|v_k - \hat{P}_{k,x^0}| \leq C_0\rho^{2+\alpha} \quad \text{in } B_\rho \cap \{x : x_n > -x_n^0/\rho^k\}$$

for $\rho < \frac{1}{4}$, which gives

$$|u_k - \hat{P}_{k,x^0}| \leq |u_k - v_k| + |v_k - \hat{P}_{k,x^0}| \leq \epsilon + C_0\rho^{2+\alpha} \quad \text{in } B_\rho \cap \{x : x_n > -x_n^0/\rho^k\}.$$

For $\rho^\alpha \leq \frac{1}{2C_0}$ and $\epsilon \leq \frac{1}{2}\rho^2$, we get

$$|u_k - \hat{P}_{k,x^0}| \leq \rho^2 \quad \text{in } B_\rho \cap \{x : x_n > -x_n^0/\rho^k\},$$

or, in other words,

$$|u - P_{k+1,x^0}| \leq \rho^{2(k+1)} \quad \text{in } B^+_{\rho^{k+1}}(x^0)$$

for

$$P_{k+1,x^0}(x) := P_{k,x^0}(x) + \rho^{2k}\hat{P}_{k,x^0}\left(\frac{x-x^0}{\rho^k}\right).$$

Also, since $F_k(D^2\hat{P}_{k,x^0}, 0) = 0$, by (10) we have

$$F(D^2 P_{k+1,x^0}, x^0) = F\big(D^2 P_{k,x^0} + D^2\hat{P}_{k,x^0}, x^0\big) = F_k(D^2\hat{P}_{k,x^0}, 0) = 0,$$
$$|D^2 P_{k+1,x^0} - D^2 P_{k,x^0}| = |D^2\hat{P}_{k,x^0}| = |D^2 v_k(0)| \leq C_0. \qquad \square$$

Proposition 3.6 (BMO-estimate). *Let u be a solution to* (9), *f bounded, and P_{k,x^0} and ρ be as in Lemma 3.5. Then*

$$\fint_{B^+_{\rho^k/2}(x^0)} |D^2u(y) - D^2P_{k,x^0}|^2 \le C, \quad x^0 \in \overline{B}^+_{1/2},$$

if ρ is smaller than a constant that depends only on $\|u\|_{W^{2,p}(B_1)}$, f, $\overline{C}$ in (H4), *and universal constants.*

Proof. Let $x^0 \in \overline{B}^+_{1/2}$ and define

$$v(x) := u\Big(\frac{x}{R}\Big) \quad \text{and} \quad G(M,x) := \frac{1}{R^2} F\Big(R^2M, \frac{x}{R}\Big)$$

for $R = R(\overline{C}, f, K, \delta)$ ($\overline{C}$ as in (H4)) chosen so that $|G(D^2v,x)| \le \delta$ in B^+_R for δ as in Lemma 3.4. Note also that

$$\beta_G(x,y) := \sup_{M \in \mathcal{S}} \frac{|G(M,x) - G(M,y)|}{|M|+1}$$

satisfies (H4). Then v solves

$$\begin{cases} G(D^2v,x) = f(x/R)/R^2 & \text{a.e. in } B^+_R \cap (R\Omega), \\ |D^2v| \le K/R^2 & \text{a.e. in } B^+_R \backslash (R\Omega), \\ v = 0 & \text{on } B'_R, \end{cases}$$

and there is a polynomial $\tilde{P}_{k,x^0}$ for which $G(D^2\tilde{P}_{k,x^0}, Rx^0) = 0$, and a constant $\tilde{\rho}$ such that

$$|v(x) - \tilde{P}_{k,x^0}(x)| \le \tilde{\rho}^{2k}, \quad x \in B^+_{\tilde{\rho}^k}(Rx^0),$$

that is,

$$|u(x) - P_{k,x^0}(x)| \le R^2\rho^{2k}, \quad x \in B^+_{\rho^k}(x^0),$$

for $P_{k,x^0}(x) := \tilde{P}_{k,x^0}(Rx)$ and $\rho^k := \tilde{\rho}^k/R$. Note also that

$$F(D^2P_{k,x^0}, x^0) = F\Big(R^2D^2\tilde{P}_{k,x^0}, \frac{Rx^0}{R}\Big) = R^2G(D^2\tilde{P}_{k,x^0}, Rx^0) = 0.$$

In particular, for

$$u_k(x) := \frac{u(\rho^kx + x^0) - P_{k,x^0}(\rho^kx + x^0)}{\rho^{2k}},$$
$$F_k(M,x) := F(M + D^2P_{k,x^0}, \rho^kx + x^0),$$

and β_k as in the proof of Lemma 3.5, we have $|u_k| \le R^2$, $\beta_k(x,y) \le \eta$ and $|F_k(u_k,x)| \le C$. Therefore we can apply Theorem 3.2 to deduce

$$\|u_k\|_{W^{2,p}(B_{1/2} \cap \{x_n \ge -x^0/\rho^k\})} \le C,$$

or

$$\fint_{B^+_{\rho^k/2}(x^0)} |D^2u(x) - D^2P_{k,x^0}|^p \, dx \le C. \qquad \square$$

From this it is straightforward to show that if u is a function satisfying (2), there exists a second-order polynomial $P_{r,x^0}(x)$ with $F(D^2 P_{r,x^0}, x^0) = f(x^0)$ such that

$$\sup_{r\in(0,1/4)} ⨍_{B_r^+(x^0)} |D^2 u(y) - D^2 P_{r,x^0}|^2\, dy \le C,$$

where $x^0 \in \overline{B}^+_{1/2}(0)$. The proof of $C^{1,1}$-regularity now follows as in [Indrei and Minne 2015] up to minor modifications (see also [Figalli and Shahgholian 2014]). The idea is that $D^2 P_{r,x^0}(x)$ provides a suitable approximation to $D^2 u(x^0)$ and one may consider two cases: first, if $D^2 P_{r,x^0}(x)$ stays bounded in r, then one can show that $D^2 u(x^0)$ is also bounded by a constant depending only on the initial ingredients; next, if $D^2 P_{r,x^0}(x)$ blows up in r, one can show that the set

$$A_r(x^0) := \frac{(B_r^+(x^0)\backslash\Omega) - x^0}{r} = B_1 \backslash \big((\Omega - x^0)/r\big) \cap \{y : y_n > -x_n^0/r\}$$

decays fast enough to ensure yet again a bound on $D^2 u(x^0)$.

Acknowledgements

We thank Henrik Shahgholian for bringing this problem to our attention. Moreover, we thank John Andersson for his help: the technique developed in this paper evolved through our interaction with him. We also thank the G. S. Magnuson Foundation for supporting this work. E. Indrei acknowledges partial support from NSF Grants OISE-0967140 (PIRE), DMS-0405343, and DMS-0635983 administered by the Center for Nonlinear Analysis at Carnegie Mellon University. Moreover, the excellent research environment provided by KTH, CMU, the Hausdorff Research Institute for Mathematics, and Universität Bonn is kindly acknowledged. Lastly, we thank Herbert Koch and an anonymous referee for providing valuable feedback on a preliminary version of this paper.

References

[Alt and Gilardi 1982] H. W. Alt and G. Gilardi, “The behavior of the free boundary for the dam problem”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **9**:4 (1982), 571–626. MR 693780 Zbl 0521.76092

[Andersson 2007] J. Andersson, “On the regularity of a free boundary near contact points with a fixed boundary”, *J. Differential Equations* **232**:1 (2007), 285–302. MR 2281197 Zbl 1106.35143

[Andersson et al. 2006] J. Andersson, N. Matevosyan, and H. Mikayelyan, “On the tangential touch between the free and the fixed boundaries for the two-phase obstacle-like problem”, *Ark. Mat.* **44**:1 (2006), 1–15. MR 2237208 Zbl 1170.35560

[Andersson et al. 2013] J. Andersson, E. Lindgren, and H. Shahgholian, “Optimal regularity for the no-sign obstacle problem”, *Comm. Pure Appl. Math.* **66**:2 (2013), 245–262. MR 2999297 Zbl 06132672

[Apushkinskaya and Uraltseva 1995] D. E. Apushkinskaya and N. N. Uraltseva, “On the behavior of the free boundary near the boundary of the domain”, *Zap. Nauchn. Sem. POMI* **221** (1995), 5–19. In Russian; translated in *J. Math. Sci. New York* **87**:2 (1997), 3267–3276. MR 1359745 Zbl 0888.35141

[Caffarelli and Cabré 1995] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, American Mathematical Society Colloquium Publications **43**, American Mathematical Society, Providence, RI, 1995. MR 1351007 Zbl 0834.35002

[Caffarelli and Gilardi 1980] L. A. Caffarelli and G. Gilardi, “Monotonicity of the free boundary in the two-dimensional dam problem”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **7**:3 (1980), 523–537. MR 597551 Zbl 0513.76090

[Caffarelli et al. 1996] L. Caffarelli, M. G. Crandall, M. Kocan, and A. Święch, "On viscosity solutions of fully nonlinear equations with measurable ingredients", *Comm. Pure Appl. Math.* **49**:4 (1996), 365–397. MR 1376656 Zbl 0854.35032

[Figalli and Shahgholian 2014] A. Figalli and H. Shahgholian, "A general class of free boundary problems for fully nonlinear elliptic equations", *Arch. Ration. Mech. Anal.* **213**:1 (2014), 269–286. MR 3198649 Zbl 1326.35137

[Indrei and Minne 2015] E. Indrei and A. Minne, "Regularity of solutions to fully nonlinear elliptic and parabolic free boundary problems", *Ann. Inst. H. Poincaré Anal. Non Linéaire* (online publication May 11 2015).

[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 661144 Zbl 0529.35026

[Matevosyan 2005] N. Matevosyan, "Tangential touch between free and fixed boundaries in a problem from superconductivity", *Comm. Partial Differential Equations* **30**:7-9 (2005), 1205–1216. MR 2180300 Zbl 1081.35161

[Matevosyan and Markowich 2004] N. Matevosyan and P. A. Markowich, "Behavior of the free boundary near contact points with the fixed boundary for nonlinear elliptic equations", *Monatsh. Math.* **142**:1-2 (2004), 17–25. MR 2065018 Zbl 1064.35215

[Petrosyan et al. 2012] A. Petrosyan, H. Shahgholian, and N. Uraltseva, *Regularity of free boundaries in obstacle-type problems*, Graduate Studies in Mathematics **136**, American Mathematical Society, Providence, RI, 2012. MR 2962060 Zbl 1254.35001

[Safonov 1994] M. Safonov, "On the boundary value problems for fully nonlinear elliptic equations of second order", Mathematics Research Report No. MRR 049-94, The Australian National University, Canberra, 1994, available at http://www.math.umn.edu/~safon002/NOTES/BVP_94/BVP.pdf.

[Shahgholian and Uraltseva 2003] H. Shahgholian and N. Uraltseva, "Regularity properties of a free boundary near contact points with the fixed boundary", *Duke Math. J.* **116**:1 (2003), 1–34. MR 1950478 Zbl 1050.35157

[Wang 1992] L. Wang, "On the regularity theory of fully nonlinear parabolic equations, II", *Comm. Pure Appl. Math.* **45**:2 (1992), 141–178. MR 1139064 Zbl 0774.35042

[Winter 2009] N. Winter, "$W^{2,p}$ and $W^{1,p}$-estimates at the boundary for solutions of fully nonlinear, uniformly elliptic equations", *Z. Anal. Anwend.* **28**:2 (2009), 129–164. MR 2486925 Zbl 1206.35116

Received 12 Jun 2015. Revised 6 Jan 2016. Accepted 9 Feb 2016.

EMANUEL INDREI: egi@cmu.edu
Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA 15213, United States

ANDREAS MINNE: minne@kth.se
Department of Mathematics, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden

ANALYSIS AND PDE
Vol. 9, No. 2, 2016
dx.doi.org/10.2140/apde.2016.9.503

CORRECTION TO THE ARTICLE SCATTERING THRESHOLD FOR THE FOCUSING NONLINEAR KLEIN–GORDON EQUATION

SLIM IBRAHIM, NADER MASMOUDI AND KENJI NAKANISHI

Volume **4**:3 (2011), 405–460

This article resolves some errors in the paper "Scattering threshold for the focusing nonlinear Klein–Gordon equation", *Anal. PDE* **4**:3 (2011), 405–460. The errors are in the energy-critical cases in two and higher dimensions.

1. The errors and the missing ingredient

This article resolves some errors in [Ibrahim et al. 2011]. One correction affects also [Ibrahim et al. 2014; 2015]; henceforth, we refer to these papers by their years only. The major errors are the following three, one in [2011, Section 2] for the existence of mass-shifted ground state in the two-dimensional energy-critical case, and two in [2011, Section 5] for the nonlinear profile decomposition in the higher-dimensional energy-critical case:

(1) In the proof of [2011, Lemma 2.6], it is not precluded that the weak limit Q in [2011, (2-67)] is zero. Hence the existence of Q in the case $c \le 1$ is not proved.

(2) In [2011, (5-56)], we do not have $\|\vec{V}_n(\tau_n) - \vec{V}_\infty(\tau_n)\|_{L^2_x} \to 0$ when $h_\infty = 0$, $\tau_\infty = \pm\infty$ and $\liminf_{n\to\infty} |\tau_n h_n^2| > 0$. Indeed, assuming that $\tau_n h_n^2 \to m \in [-\infty, \infty]$ after extraction of a subsequence, we have

$$\|\vec{V}_n(\tau_n) - \vec{V}_\infty(\tau_n)\|_{L^2_x} \to \begin{cases} \|(e^{im/(2|\nabla|)} - 1)\psi\|_{L^2_x} & (|m| < \infty), \\ \sqrt{2}\|\psi\|_{L^2_x} & (m = \pm\infty). \end{cases} \tag{1-1}$$

(3) In the proof of [2011, Lemma 5.6], the global bound [2011, (5-96)] does not follow from the uniform bound on finite time intervals, since the required largeness of n depends on the size of the interval I.

(1) is concerned only with a very critical case of exponential nonlinearity in two dimensions ($d = 2$). More precisely, it is problematic only if

$$0 < \limsup_{|u|\to\infty} e^{-\kappa_0|u|^2}|u|^2 f(u) < \infty, \tag{1-2}$$

MSC2010: 35B40, 35L70, 47J30, 35B44.
Keywords: nonlinear Klein–Gordon equation, scattering theory, blow-up solution, ground state, Sobolev critical exponent, Trudinger–Moser inequality.

where κ_0 is the exponent in [2011, (1-29)]. Errors (2)–(3) are crucial only in the H^1 critical case of higher dimensions $d \geq 3$, with $h_\infty = 0$: the concentration by scaling in the nonlinear profile, where we need to modify the definition of the nonlinear concentrating waves and then solve the massless limit problem for the nonlinear Klein–Gordon equation (NLKG) (see Theorem 3.1 below). In the other case, i.e., with the subcritical or exponential nonlinearity or with $h_\infty = 1$, we still need to take care of (3), but it is a rather superficial change.

2. Correction for (1)

We do not know if [2011, Lemma 2.6] holds true in the very critical case (1-2). So we add the assumption

$$\limsup_{|u|\to\infty} e^{-\kappa_0|u|^2}|u|^2 f(u) \in \{0, \infty\} \tag{2-1}$$

in [2011, Proposition 1.2(3)] and [2011, Lemma 2.6]. The existence of Q was used in [2011] only to characterize the threshold energy m, so the rest of the paper is not affected by it.

In [2014, (1.24)], the existence of Q is mentioned to characterize the threshold $m^{(c)}$. It should also be restricted by (2-1), but the rest of [2014] does not really need Q. Removing Q, [2014, (2.3)] should be replaced with

$$m \leq H_p^{(c)}(\varphi), \tag{2-2}$$

[2014, (2.6)] should be replaced with

$$m \leq J^{(c)}(\lambda\varphi) = H_p^{(c)}(\lambda\varphi) \leq H_p^{(c)}(\varphi), \tag{2-3}$$

and [2014, (2.7)] with

$$\begin{aligned}\ddot{y} = (2+p)\|\dot{u}\|_{L^2}^2 + 2p(H_p^{(1)}(u) - m) &= (4+\varepsilon)\|\dot{u}\|_{L^2}^2 + (1-c)\varepsilon\|u\|_{L^2}^2 + 2p(H_p^{(c)}(u) - m) \\ &\geq \left(1 + \tfrac{1}{4}\varepsilon\right)\dot{y}^2/y + (1-c)\varepsilon y.\end{aligned} \tag{2-4}$$

The existence of Q is also mentioned in [2015, Theorem 5.1]. It should be also restricted by (2-1). The rest of [2015] remains unaffected.

We still need to prove [2011, Lemma 2.6] under the new restriction (2-1). If the limit (2-1) is infinite, then [2015, Theorem 1.5(B)] implies $C^\star_{\mathrm{TM}}(F) = \infty > 1$. In this case, the proof of [2011, Lemma 2.6] remains valid. If the limit (2-1) is zero, then [2015, Theorem 1.5(B)] implies $C^\star_{\mathrm{TM}}(F) < \infty$. In this case, we do not argue as in [2011], but rely on the compactness [2015, Theorem 1.5(C)]. Let $\varphi_n \in H^1(\mathbb{R}^2)$ be a normalized maximizing sequence for $C^\star_{\mathrm{TM}}(F)$, i.e.,

$$\|\varphi_n\|_{L^2} = 1, \quad \kappa_0\|\nabla\varphi_n\|_{L^2}^2 \leq 4\pi, \quad 2F(\varphi_n) \to C := C^\star_{\mathrm{TM}}(F) \in (0, \infty). \tag{2-5}$$

By the standard rearrangement and the H^1 boundedness, we may assume that the φ_n are radially decreasing and $\varphi_n \to \varphi$ weakly in $H^1(\mathbb{R}^2)$ for some φ. By [2015, Theorem 1.5(C)], we have $2F(\varphi_n) \to 2F(\varphi) = C > 0$. In particular, $\varphi \neq 0$. Since $\kappa_0\|\nabla\varphi\|_{L^2}^2 \leq 4\pi$ and $\|\varphi\|_{L^2} \leq 1$ by the weak convergence, we deduce from the definition of $C^\star_{\mathrm{TM}}(F)$ that $\|\varphi\|_{L^2} = 1$ and φ is a maximizer. Hence, for a Lagrange multiplier $\mu \geq 0$,

$$f'(\varphi) - C\varphi = -\mu\Delta\varphi. \tag{2-6}$$

That $\mu \neq 0$ is obvious by the decay order of f' as $\varphi \to 0$. Hence $\mu > 0$ and so $\kappa_0 \|\nabla\varphi\|_{L^2}^2 = 4\pi$, since otherwise we could increase both $F(\varphi)$ and $\|\nabla\varphi\|_{L^2}^2$ by the L^2 scaling $\varphi_{1,-1}^{\lambda}$ with $\lambda > 0$, using the L^2 supercritical condition [2011, (1-21)]. Then $Q(x) := \varphi(\mu^{-1/2}x) \in H^2(\mathbb{R}^2)$ satisfies

$$-\Delta Q + CQ = f'(Q), \quad \kappa_0\|\nabla Q\|_{L^2}^2 = 4\pi, \quad 2F(Q) = C\|Q\|_{L^2}^2. \tag{2-7}$$

Hence $J^{(C)}(Q) = \frac{1}{2}\|\nabla Q\|_{L^2}^2 = 2\pi/\kappa_0$. The rest of the proof of [2011, Lemma 2.6], namely the proof of $m_{\alpha,\beta} = m_{0,1} = 2\pi/\kappa_0$, remains valid.

3. Correction for (2)–(3)

For (2)–(3), we do not have to modify the main results, but need to correct the proof, including the definition of the nonlinear profile decomposition. Henceforth, we always assume that $0 < h_n \to h_\infty$, $(t_n, x_n) \in \mathbb{R}^{1+d}$ and $\tau_n = -t_n/h_n \to \tau_\infty \in [-\infty, \infty]$ are sequences. The main problematic case is when the energy concentrates, namely $h_\infty = 0$, which can happen only in the energy critical case [2011, (1-28)]

$$d \geq 3, \quad f(u) = \frac{|u|^{2^\star}}{2^\star}, \quad 2^\star = \frac{2d}{d-2}. \tag{3-1}$$

First we modify the vector notation in [2011, (4-1)]. For any real-valued function $a(t,x)$, the complex-valued functions $\vec{a}$, $\vec{a}$ and $\overset{\rightharpoonup}{a}$ are defined by

$$\vec{a} := (\langle\nabla\rangle - i\partial_t)a, \quad \vec{a} := (\langle\nabla\rangle_n - i\partial_t)a, \quad \overset{\rightharpoonup}{a} := (\langle\nabla\rangle_\infty - i\partial_t)a, \tag{3-2}$$

where $\langle\nabla\rangle_* = \sqrt{h_*^2 - \Delta}$ as in [2011, (5-1)]. Hence a is recovered from either of them by

$$a = \operatorname{Re}\langle\nabla\rangle^{-1}\vec{a} = \operatorname{Re}\langle\nabla\rangle_n^{-1}\vec{a} = \operatorname{Re}\langle\nabla\rangle_\infty^{-1}\overset{\rightharpoonup}{a}. \tag{3-3}$$

Note that $(\vec{a}, a)$ was denoted by $(\vec{a}, \hat{a})$ in [2011], but it was confusing. Indeed, $u_{(n)}$ in [2011, (5-55)] did not make sense if $h_\infty = 0$, since $\vec{u}_{(n)}$ in [2011, (5-54)] was not in the form [2011, (4-1)]. So we replace [2011, (5-54)] with

$$\vec{u}_{(n)} = T_n \vec{U}_{(n)}((t-t_n)/h_n), \tag{3-4}$$

where $\vec{U}_{(n)}$ is defined by

$$\vec{V}_n := e^{it\langle\nabla\rangle_n}\psi, \quad \vec{U}_{(n)} = \vec{V}_n - i\int_{\tau_\infty}^{t} e^{i(t-s)\langle\nabla\rangle_n} f'(U_{(n)})\,ds. \tag{3-5}$$

Then $u_{(n)} = h_n T_n U_{(n)}((t-t_n)/h_n)$ is a solution of NLKG satisfying

$$\lim_{t\to\tau_\infty} \|(\vec{u}_{(n)} - \vec{v}_n)(th_n + t_n)\|_{L_x^2} = 0. \tag{3-6}$$

In other words, we keep NLKG in defining the profiles, even if $h_\infty = 0$. Note that if $h_\infty = 1$ then $\vec{U}_{(n)} = \vec{U}_\infty$ and so $u_{(n)}$ is unchanged.

By the change of [2011, (5-54)] to (3-4), the problematic [2011, (5-56)] is replaced with

$$\|\vec{u}_n(0) - \vec{u}_{(n)}(0)\|_{L_x^2} = \left\| \int_{\tau_\infty h_n + t_n}^{0\ (=\tau_n h_n + t_n)} e^{-is\langle\nabla\rangle} f'(u_{(n)})\,ds \right\|_{L_x^2} \to 0. \tag{3-7}$$

In order to prove the last limit as well as the global Strichartz approximation for (3), we need the convergence in the massless limit of the H^1 critical NLKG:

Theorem 3.1. *Assume* [2011, (1-28)] *and $h_\infty = 0$. Let $\vec{U}_\infty$ be the solution of*

$$\vec{V}_\infty := e^{it|\nabla|}\psi, \quad \vec{U}_\infty = \vec{V}_\infty - i\int_{\tau_\infty}^{t} e^{i(t-s)|\nabla|} f'(U_\infty)\, ds. \tag{3-8}$$

Let $\vec{U}_{(n)}$ be the solution of (3-5) *and $\vec{u}_{(n)}(t) := T_n\vec{U}_{(n)}((t-t_n)/h_n)$. Suppose that $U_\infty \in [W]_2^{\bullet}(J)$ for some interval J whose closure in* $[-\infty, \infty]$ *contains τ_∞. Then for any bounded subinterval $I \subset J$ we have, as $n \to \infty$,*

$$\begin{gathered}\|\vec{U}_{(n)} - \vec{U}_\infty\|_{L^\infty_{t\in I}L^2_x} + \|U_{(n)} - U_\infty\|_{([W]_2^\bullet\cap[M]_0)(J)} + \|u_{(n)}\|_{[W]_0(J)} \to 0,\\ \|u_{(n)}\|_{([W]_2\cap[M]_0)(h_nJ+t_n)} \sim \|U_\infty\|_{([W]_2^\bullet\cap[M]_0)(J)} + o(1).\end{gathered} \tag{3-9}$$

Postponing the proof of the above theorem to the next section, we continue to correct [2011, Section 5]. Equation (3-7) in the case of $h_\infty = 0$ follows from the above estimate and $\tau_n \to \tau_\infty$ via Strichartz:

$$\begin{aligned}\left\|\int_{\tau_\infty h_n+t_n}^{0} e^{-is\langle\nabla\rangle} f'(u_{(n)})\, ds\right\|_{L^2_x} &\lesssim \|f'(u_{(n)})\|_{[W^{*(1)}]_2(I_n)} \lesssim \|u_{(n)}\|^{2^\star-1}_{([W]_2\cap[M]_0)(I_n)}\\ &\lesssim \|U_\infty\|^{2^\star-1}_{[W]_2^\bullet\cap[M]_0(J_n)} + o(1) = o(1),\end{aligned} \tag{3-10}$$

where $I_n := (0, \tau_\infty h_n + t_n) \cup (\tau_\infty h_n + t_n, 0)$ and $J_n := (\tau_n, \tau_\infty) \cup (\tau_\infty, \tau_n)$.

We modify the definition of ST in [2011, (5-59)–(5-60)] in the $\dot{H}^1$ critical case [2011, (1-28)] to

$$ST = [W]_2, \quad ST^* = [W^{*(1)}]_2 + L^1_t L^2_x, \quad ST^\diamond_\infty := \begin{cases}[W]_2 & (h^\diamond_\infty = 1),\\ [W]_2^\bullet & (h^\diamond_\infty = 0).\end{cases} \tag{3-11}$$

Indeed, $[K]_2$ and $[K^{*(1)}]_2$ norms are not needed in the $\dot{H}^1$ critical case. Then we simply discard the estimates [2011, (5-61)–(5-62)].

Next we reprove [2011, Lemma 5.5], extending it to unbounded intervals I. The above theorem implies that we can replace [2011, (5-64)] with the stronger[1]

$$\limsup_{n\to\infty} \|u^j_{(n)}\|_{ST(\mathbb{R})} \lesssim \|U^j_\infty\|_{ST^j_\infty(\mathbb{R})} \tag{3-12}$$

if $h^j_\infty = 0$, while it is trivial if $h^j_\infty = 1$. The proof of [2011, (5-65)] for $h^j_\infty = 1$ did not use the boundedness of I, so we may assume that all h^j_∞ are 0. Then the above theorem implies that $\|u^{<k}_{(n)}\|_{[W]_0(\mathbb{R})} \to 0$ as $n \to \infty$, so it suffices to estimate the homogeneous norm $[W]_2^\bullet(\mathbb{R})$. We have

$$\|u^{<k}_{(n)}\|_{[W]_2^\bullet(\mathbb{R})} \sim \sum_{l=1}^{d} \left\| \sum_{j<k} \check{u}^{j,l}_{n,m} \right\|_{L^p_t \ell^2_{m\in\mathbb{Z}} L^q_x} \tag{3-13}$$

with $(1/p, 1/q, s) = W$ and

$$\check{u}^{j,l}_{n,m} := 2^{sm}\delta^l_m h^j_n T^j_n U^j_{(n)}((t-t^j_n)/h^j_n). \tag{3-14}$$

[1]Recall that $\widehat{U}^j_\infty$ in [2011] is denoted by U^j_∞ in this correction according to (3-2).

Defining $\check{u}^{j,l}_{n,m,R}$ by [2011, (5-77)], we have

$$\|\check{u}^{j,l}_{n,m} - \check{u}^{j,l}_{n,m,R}\|_{L^p_t \ell^2_m L^q_x} \lesssim \|2^{sm}\delta^l_m U^j_{(n)}\|_{L^p_t \ell^2_m L^q_x(|t|+|m|+|x|>R)} \to 0 \quad \text{as } R \to \infty, \tag{3-15}$$

which is still uniform in n since, by the above theorem, $U^j_{(n)}$ is approximated by U^j_∞ in $[W]^\bullet_2(\mathbb{R})$, which is equivalent to the last norm without the restriction by R. Thus we obtain [2011, (5-65)] by the disjoint support property for large n.

According to the change of $u^j_{(n)}$, we replace the nonlinear decomposition [2011, (5-66)] with the simpler form

$$\lim_{n\to\infty} \Big\| f'(u^{<k}_{(n)}) - \sum_{j<k} f'(u^j_{(n)}) \Big\|_{ST^*(I)} = 0, \tag{3-16}$$

which is the same as [2011, (5-66)] if $h^j_\infty = 1$. In that case, however, we used that I was bounded in [2011, (5-82)]. We replace it with an interpolation between [2011, (4-84)] and

$$\|f'_S(u)\|_{[((1-\theta_0)K+\theta_0 W)^{*(1)}]_2(I)} \lesssim \|u\|_{[K]_2(I)} \|u\|^{p_1}_{[K]_0(I)} \lesssim \|u\|^{p_1+1}_{[K]_2(I)}, \tag{3-17}$$

where we can choose some $\theta_0 \in (0,1)$ since $p_1 > 4/d$ (choosing p_1 close enough to $4/d$ if necessary). Since $Z := ((1-\theta_0)K + \theta_0 W)^{*(1)}$ is an interior dual-admissible exponent, we can find some $\theta_1 \in (0,1)$ such that $\theta_1 Y + (1-\theta_1)Z$ is also a dual-admissible exponent. Interpolating (3-17) with [2011, (4-84)], we have

$$\|f'_S(u) - f'_S(v)\|_{[\theta_1 Y + (1-\theta_1)Z]_2(I)} \lesssim \|(u,v)\|^{p_1+1-\theta_1}_{[K]_2(I)\cap[Q]_{2p_1}(I)} \|u-v\|^{\theta_1}_{[P]_2(I)}. \tag{3-18}$$

Thus we obtain [2011, (5-66)] on any subset I in the subcritical and exponential cases. In the $\dot{H}^1$ critical case [2011, (1-28)], we discard $u^j_{\langle n\rangle}$ in [2011, (5-85)] and prove (3-16) directly, putting

$$U^j_{n,R}(t,x) := \chi_R(t,x) U^j_{(n)}(t,x) \times \prod \big\{ (1-\chi_{h^{j,l}_n R})(t - t^{j,l}_n, x - x^{j,l}_n) \;\big|\; 1 \le l < k,\ h^l_n R < h^j_n \big\}. \tag{3-19}$$

It is still uniformly bounded in $([H]^\bullet_2 \cap [W]^\bullet_2)(\mathbb{R})$, and $U^j_{n,R} - \chi_R U^j_{(n)} \to 0$ in $[M]_0(\mathbb{R})$ as $n \to \infty$ thanks to the above theorem, as well as in $[L]_0$, and also $\chi_R U^j_{(n)} \to U^j_{(n)}$ as $R \to \infty$. Hence we may replace $u^j_{(n)}$ in (3-16) by $u^j_{(n),R} := h^j_n T^j_n U^j_{n,R}((t - t^j_n)/h^j_n)$, using [2011, (4-62)] for $d \le 5$, and a similar interpolation argument as above for $d \ge 6$; see (4-16)–(4-19) below. Then we obtain (3-16) by the disjoint support property, in the same way as [2011, (5-94)].

With the above corrections, we now reprove [2011, Lemma 5.6]. First, [2011, (5-100)] holds for any subset $I \subset \mathbb{R}$, by the above improvement of [2011, Lemma 5.5]. Now, thanks to the change of $u^j_{(n)}$, [2011, (5-101)] is simplified to

$$\mathrm{eq}(u^{<k}_{(n)}) = f'(u^{<k}_{(n)}) - \sum_{j<k} f'(u^j_{(n)}), \tag{3-20}$$

which is vanishing by (3-16). Hence we obtain [2011, (5-103)]. We also obtain [2011, (5-104)] on $\mathbb{R}$ by the same nonlinear estimates as we used above. Then, applying [2011, Lemma 4.5] on $\mathbb{R}$, we obtain the desired [2011, Lemma 5.6].

Section 6 of [2011] is almost unchanged, except for the obvious modification in [2011, (6-6)] due to the change of $u_{(n)}$, namely

$$\vec{u}^{j}_{(n)} = T^j_n \vec{U}^j_{(n)}((t-t^j_n)/h^j_n), \tag{3-21}$$

and the notational change in [2011, (6-7)–(6-9)] from $(\vec{U}^0_\infty, \widehat{U}^0_\infty)$ to $(\vec{U}^0_\infty, U^0_\infty)$ due to (3-2). Since the case $h_\infty = 0$ is eliminated in the proof of [2011, Lemma 6.1], the errors (2)–(3) do not affect the rest of the paper.

4. Massless limit of scattering for the critical NLKG

It remains to prove Theorem 3.1. Throughout this section, we assume [2011, (1-28)]. The main idea is to decompose the time interval into a bounded subinterval and neighborhoods of $\pm\infty$. On the bounded part, we have strong convergence in the massless limit. In the neighborhoods of $t = \pm\infty$, we do not have strong convergence, but the Strichartz norms are uniformly controlled via the asymptotic free profiles.

The first ingredient concerns the uniform Strichartz bound for free waves.

Lemma 4.1. *Let* $\vec{v}_n = e^{it\langle\nabla\rangle} T_n \psi$, $h_\infty = 0$, $\vec{V}_\infty = e^{it|\nabla|}\psi$, *and let* $Z \in \left[0, \tfrac{1}{2}\right] \times \left[0, \tfrac{1}{2}\right) \times [0, 1)$ *satisfy* $\mathrm{reg}^0(Z) = 1$ *and* $\mathrm{str}^0(Z) \le 0$, *namely a wave-admissible Strichartz exponent except for the energy norm. Then we have*

$$\limsup_{n\to\infty} \|v_n\|_{[Z]_2(0,\infty)} \lesssim \|V_\infty\|_{[Z]^\bullet_2(0,\infty)} \quad \text{and} \quad \lim_{n\to\infty} \|P_{<1} v_n\|_{[Z]_2(0,\infty)} = 0, \tag{4-1}$$

where $P_{<a}$ *denotes the smooth cut-off for the Fourier region* $|\xi| < 2a$ *defined by* $P_{<a}\varphi = a^d \Lambda_0(ax) * \varphi$, *with* $\Lambda_0 \in \mathcal{S}(\mathbb{R}^d)$ *in the proof of* [2011, Lemma 5.1]. *If* $Z_3 = 0$, *then we have also* $\|v_n\|_{[Z]_0(0,\infty)} \to \|V_\infty\|_{[Z]_0(0,\infty)}$.

Proof. Let $\vec{v}_n(t) = T_n \vec{V}_n(t/h_n)$. The Strichartz estimate for the Klein–Gordon and the wave equations

$$\|v_n\|_{[Z]_2(0,\infty)} \lesssim \|T_n\psi\|_{L^2} = \|\psi\|_{L^2}, \quad \|V_\infty\|_{[Z]^\bullet_2(0,\infty)} \lesssim \|\psi\|_{L^2} \tag{4-2}$$

implies that it suffices to consider ψ in a dense subset of $L^2(\mathbb{R}^d)$. Hence we may assume that $\mathcal{F}\psi$ is C^∞ with a compact $\operatorname{supp}\mathcal{F}\psi \not\ni 0$. Since $0 < \langle\xi\rangle_n - \langle\xi\rangle_\infty \le h_n^2/|\xi|$,

$$|(e^{it\langle\xi\rangle_n}\langle\xi\rangle_n^{-1} - e^{it|\xi|}|\xi|^{-1})| \lesssim |t| h_n^2 |\xi|^{-2} + h_n^2|\xi|^{-3}, \tag{4-3}$$

and so, under the above assumption on ψ, for any $s \in \mathbb{R}$ and any sequence $S_n > 0$,

$$\|V_n - V_\infty\|_{L^\infty(0,S_n;H^s)} \le \langle S_n\rangle h_n^2 C(s,\psi). \tag{4-4}$$

Hence, by Sobolev in x and Hölder in t,

$$\|V_n - V_\infty\|_{([Z]^\bullet_2 \cap [Z]_0)(0,S_n)} \le \langle S_n\rangle^{1+Z_1} h_n^2 C(s,\psi). \tag{4-5}$$

We deduce that if $S_n \to \infty$ and $S_n^{1+Z_1} h_n^2 \to 0$ then, using the (approximate) scale-invariance of $[Z]^\bullet_2$,

$$\begin{aligned} \|v_n\|_{[Z]_2(0,h_nS_n)} &\sim \|v_n\|_{[Z]^\bullet_2(0,h_nS_n)} + \|P_{<1}v_n\|_{[Z]_0(0,h_nS_n)}, \\ \|v_n\|_{[Z]^\bullet_2(0,h_nS_n)} &\sim \|V_n\|_{[Z]^\bullet_2(0,S_n)} \to \|V_\infty\|_{[Z]^\bullet_2(0,\infty)}, \\ \|P_{<1}v_n\|_{[Z]_0(0,h_nS_n)} &\sim \|h_n^{Z_3} P_{<h_n} V_n\|_{[Z]_0(0,S_n)} \to 0, \end{aligned}$$

and similarly, if $Z_3 = 0$ then $\|v_n\|_{[Z]_0(0,h_nS_n)} = \|V_n\|_{[Z]_0(0,S_n)} \to \|V_\infty\|_{[Z]_0(0,\infty)}$.

Next, the dispersive decay of wave-type for the Klein–Gordon equation

$$\|e^{it\langle\nabla\rangle}\varphi\|_{B^0_{q,2}} \lesssim |t|^{-(d-1)\alpha}\|\varphi\|_{B^s_{q',2}}, \quad \alpha := \tfrac{1}{2} - \tfrac{1}{q} \in \big[0, \tfrac{1}{2}\big],\ s := (d+1)\alpha, \tag{4-6}$$

together with the embedding $L^{q'} \subset B^0_{q',2}$ implies that

$$\|v_n(t)\|_{B^\sigma_{q,2}} \lesssim |t|^{-(d-1)\alpha}\|\langle\nabla\rangle^{\sigma+s-1}T_n\psi\|_{L^{q'}} = |t|^{-(d-1)\alpha}h_n^{1-\alpha-\sigma}\|\langle\nabla\rangle_n^{\sigma+s-1}\psi\|_{L^{q'}}, \tag{4-7}$$

and so, putting $\alpha = \frac{1}{2} - Z_2$,

$$\begin{aligned}\|v_n\|_{[Z]_2(h_nS_n,\infty)} &\le C(\psi)h_n^{1-\alpha-Z_3}\|t^{-(d-1)\alpha}\|_{L_t^{1/Z_1}(h_nS_n,\infty)}\\ &\sim C(\psi)h_n^{1-\alpha-Z_3}(h_nS_n)^{Z_1-(d-1)\alpha} = C(\psi)S_n^{\alpha-1+Z_3} \to 0,\end{aligned} \tag{4-8}$$

where we used that $\mathrm{reg}^0(Z) = Z_3 - Z_1 + d\alpha = 1$ in the last identity and

$$\alpha - 1 + Z_3 = \mathrm{reg}^0(Z) + \mathrm{str}^0(Z) - 1 - Z_1 < 0 \tag{4-9}$$

in taking the limit. Note that the above exponent is zero at the energy space $Z = \big(0, \frac{1}{2}, 1\big)$, which is excluded by the assumption. The estimate in $[Z]_0(h_nS_n,\infty)$ for $Z_3 = 0$ is done in the same way. Combining them with the above estimates on $(0, h_nS_n)$ leads to the conclusion via a density argument. □

The second ingredient is convergence or propagation of small disturbance on finite intervals, which is uniformly controlled by the Strichartz norm of U_∞.

Lemma 4.2. *For any $0 < M, \varepsilon < \infty$, there exists $\delta = \delta(\varepsilon, M) \in (0, 1)$ with the following property. Let $h_\infty = 0$ and let U_∞ be a solution of NLW on some interval J satisfying $\|U_\infty\|_{([H]^\bullet_2\cap[W]^\bullet_2)(J)} \le M$. Then, for any bounded subinterval $I \subset J$ with $0 \in I$ and any $\varphi_n \in L^2(\mathbb{R}^d)$ with $\|\varphi_n\|_{L^2} < \delta$, the unique solution U_n of*

$$(\partial_t^2 - \Delta + h_n^2)U_n = f'(U_n), \quad \vec{U}_n(0) = \vec{U}_\infty(0) + \varphi_n, \tag{4-10}$$

exists on I for large n, satisfying

$$\|\vec{U}_n - \vec{U}_\infty\|_{L_t^\infty L_x^2(I)} + \|U_n - U_\infty\|_{([W]^\bullet_2\cap[M]_0)(I)} < \varepsilon, \tag{4-11}$$

and $\|h_nT_nU_n((t-t_n)/h_n)\|_{[W]_0(h_nI+t_n)} \lesssim \delta$ for large n.

Proof. We give the detail only in the harder case $d \ge 6$, where we need the exotic Strichartz norms. Let $\gamma_n := U_n - U_\infty$ and $\vec{\gamma}_n := \vec{U}_n - \vec{U}_\infty$, then

$$(\partial_t^2 - \Delta)\gamma_n = f'(U_\infty + \gamma_n) - f'(U_\infty) - h_n^2U_n. \tag{4-12}$$

Note however that $\vec{\gamma}_n$ is not written only by γ_n. It suffices to prove the following:

Claim. There exist constants $\theta \in (0, 1)$ and $C > 1$ such that if

$$\|U_\infty\|_{([W]^\bullet_2\cap[\widetilde{M}]^\bullet_{2p})(0,S)} \le \eta, \quad \|\vec{\gamma}_n(0)\|_{L^2} \ll 1, \tag{4-13}$$

for some $0 < S < \infty$ and $0 < \eta \ll 1$, where $p = 2^\star - 2 = 4/(d-2)$, then

$$\|\vec{\gamma}_n\|_{L_t^\infty(0,S;L_x^2)} + \|\gamma_n\|_{[W]_2^\bullet(0,S)} \le C[\|\vec{\gamma}_n(0)\|_{L^2} + \|\vec{\gamma}_n(0)\|_{L^2}^\theta \eta^{(p+1)(1-\theta)}]. \tag{4-14}$$

Proof of the claim. The exotic Strichartz estimate for the wave equation yields, on the time interval $(0, S)$,

$$\|\gamma_n\|_{[\widetilde{N}]_2^\bullet} \lesssim \|\bar{\gamma}_n(0)\|_{L^2} + \|f'(U_\infty+\gamma_n) - f'(U_\infty)\|_{[Y]_2} + \|h_n^2 U_n\|_{L_t^1 L_x^2}, \tag{4-15}$$

while the nonlinear estimate in the Besov space yields

$$\|f'(U_\infty+\gamma_n) - f'(U_\infty)\|_{[Y]_2} \lesssim \|(U_\infty,\gamma_n)\|_{[M]_0}^p \|\gamma_n\|_{[\widetilde{N}]_2^\bullet} + \|(U_\infty,\gamma_n)\|_{[\widetilde{M}]_{2p}^\bullet}^p \|\gamma_n\|_{[N]_0}, \tag{4-16}$$

and we have $\|\bar{\gamma}_n(0)\|_{L^2} \lesssim \|\vec{\gamma}_n(0)\|_{L^2} + o(1)$. The $L_t^1 L_x^2$ norm is estimated by

$$\|h_n^2 U_n\|_{L_t^1 L_x^2} \le \|h_n \vec{U}_n\|_{L_t^1 L_x^2} \le h_n S \|\vec{\gamma}_n + \vec{U}_\infty\|_{L_t^\infty L_x^2}. \tag{4-17}$$

Define $\underline{W}, O \in \left[0, \tfrac{1}{2}\right]^3$ by

$$\begin{aligned} \underline{W} &:= W - \frac{1}{2}\Big(0, \frac{1}{d}, 1\Big) = \Big(\frac{d-1}{2(d+1)}, \frac{d^2-2d-1}{2d(d+1)}, 0\Big), \\ O &:= W + p\underline{W} = \Big(\frac{(d+2)(d-1)}{2(d+1)(d-2)}, \frac{d^3+d^2-6d-4}{2(d-2)d(d+1)}, \frac{1}{2}\Big). \end{aligned} \tag{4-18}$$

Then O is an interior dual exponent of the standard Strichartz, and so there is small $\theta \in (0, 1)$ such that $\theta Y + (1-\theta)O$ is also a dual exponent. Hence the standard Strichartz yields, for any wave-admissible exponent Z,

$$\|\gamma_n\|_{[Z]_2^\bullet} + \|\bar{\gamma}_n\|_{L_t^\infty L_x^2} \lesssim \|\bar{\gamma}_n(0)\|_{L^2} + \|f'(U_\infty+\gamma_n) - f'(U_\infty)\|_{[\theta Y+(1-\theta)O]_2^\bullet} + \|h_n^2 U_n\|_{L_t^1 L_x^2}, \tag{4-19}$$

where the nonlinear part is already estimated in $[Y]_2^\bullet$, while

$$\|f'(U_\infty+\gamma_n)\|_{[O]_2^\bullet} + \|f'(U_\infty)\|_{[O]_2^\bullet} \lesssim \eta^{p+1} + \|\gamma_n\|_{[W]_2^\bullet}^{p+1}. \tag{4-20}$$

Hence we have

$$\begin{aligned} \|\gamma_n\|_{[\widetilde{N}]_2^\bullet} &\lesssim \|\bar{\gamma}_n(0)\|_{L^2} + A + B, \\ \|\gamma_n\|_{[W]_2^\bullet \cap [\widetilde{M}]_{2p}^\bullet} + \|\bar{\gamma}_n\|_{L_t^\infty L_x^2} &\lesssim \|\bar{\gamma}_n(0)\|_{L^2} + A^\theta(\eta + \|\gamma_n\|_{[W]_2^\bullet})^{(1-\theta)(p+1)} + B, \\ A &\lesssim (\eta + \|\gamma_n\|_{[\widetilde{M}]_{2p}^\bullet})^p \|\gamma_n\|_{[\widetilde{N}]_2^\bullet}, \\ B &\lesssim S h_n \|\vec{\gamma}_n\|_{L_t^\infty L_x^2} + o(1). \end{aligned} \tag{4-21}$$

Assuming that $\|\gamma_n\|_{[\widetilde{M}]_{2p}^\bullet} \ll 1$ and that $\|\vec{\gamma}_n\|_{L_t^\infty L_x^2}$ is bounded in n, we deduce from the above estimates that

$$\begin{aligned} A \ll \|\gamma_n\|_{[\widetilde{N}]_2^\bullet} &\lesssim \|\bar{\gamma}_n(0)\|_{L^2} + o(1), \quad B = o(1), \\ \|\gamma_n\|_{[W]_2^\bullet \cap [\widetilde{M}]_{2p}^\bullet} + \|\bar{\gamma}_n\|_{L_t^\infty L_x^2} &\lesssim \|\bar{\gamma}_n(0)\|_{L^2} + \|\bar{\gamma}_n(0)\|_{L^2}^\theta \eta^{(1-\theta)(p+1)} + o(1). \end{aligned} \tag{4-22}$$

It remains to prove the uniform bound on $\|\vec{\gamma}_n\|_{L_t^\infty L_x^2}$. Let V_∞, V_n, v_n be the free solutions defined by

$$\vec{V}_\infty := e^{it|\nabla|}\vec{U}_\infty(0), \quad \vec{V}_n := e^{it\langle\nabla\rangle_n}\vec{U}_n(0), \quad \vec{v}_n = T_n \vec{V}_n(t/h_n). \tag{4-23}$$

For any $0 < R_n \to 0$ such that $h_n/R_n \to 0$, we have

$$\|\mathcal{F}\vec{\gamma}_n\|_{L^\infty(0,S;L^2(|\xi|>R_n))} \lesssim \|\overline{\gamma}_n\|_{L^\infty(0,S;L^2_x)} + o(1). \tag{4-24}$$

For the lower frequency, we have, by the energy inequality, Hölder and Sobolev,

$$\begin{aligned}\|\vec{U}_n - \vec{V}_n\|_{L^\infty_t \dot{H}^{-1}_x(0,S)} &\lesssim \|f'(U_n)\|_{L^1_t \dot{H}^{-1}_x(0,S)} \\ &\lesssim S\|U_n\|^{p+1}_{L^\infty_t \dot{H}^1_x(0,S)} \lesssim S(\|\vec{U}_\infty\|_{L^\infty_t L^2_x(0,S)} + \|\overline{\gamma}_n\|_{L^\infty_t L^2_x(0,S)})^{p+1},\end{aligned} \tag{4-25}$$

and similarly $\|\vec{U}_\infty - \vec{V}_\infty\|_{L^\infty_t \dot{H}^{-1}_x(0,S)} \lesssim S\|\vec{U}_\infty\|^{p+1}_{L^\infty_t L^2_x}$. Since $|\langle\xi\rangle_n - \langle\xi\rangle_\infty| \le h_n$, we have also $\|\vec{V}_n(t) - \vec{V}_\infty(t)\|_{L^2_x} \lesssim |t|h_n\|\vec{U}_\infty(0)\|_{L^2} + \delta$. Hence

$$\begin{aligned}&\|\mathcal{F}\vec{\gamma}_n\|_{L^\infty(0,S;L^2(|\xi|<R_n))} \\ &\quad\le R_n\|\vec{U}_n - \vec{V}_n\|_{L^\infty_t \dot{H}^{-1}_x(0,S)} + \|\vec{V}_n - \vec{V}_\infty\|_{L^\infty_t L^2_x(0,S)} + R_n\|\vec{V}_\infty - \vec{U}_\infty\|_{L^\infty_t \dot{H}^{-1}_x(0,S)} \\ &\quad\lesssim o(1)S\|\overline{\gamma}_n\|^{p+1}_{L^\infty_t L^2_x(0,S)} + \delta + o(1)\end{aligned} \tag{4-26}$$

Adding this to (4-24), we obtain

$$\|\vec{\gamma}_n\|_{L^\infty_t L^2_x(0,S)} \lesssim \|\overline{\gamma}_n\|_{L^\infty_t L^2_x(0,S)} + o(1)S\|\overline{\gamma}_n\|^{p+1}_{L^\infty_t L^2_x(0,S)} + \delta + o(1). \tag{4-27}$$

Combining this with the estimates (4-22), we deduce that both $\vec{\gamma}_n$ and $\overline{\gamma}_n$ are bounded in $L^\infty_t L^2_x(0, S)$. □

To prove (4-11) from this claim, we decompose I into subintervals I_j such that $\|U_\infty\|_{([W]^\bullet_2\cap[\widetilde{M}]^\bullet_{2p})(I_j)} \le \eta$ for each j. Then applying the above claim iteratively to the subintervals for small $\delta > 0$ yields (4-11), where the bound on $[M]_0$ is derived by interpolation and Sobolev embedding of $[H]^\bullet_2$ and $[W]^\bullet_2$.

For the estimate in $[W]_0$, we have, by scaling,

$$\begin{aligned}&\|h_nT_nU_n((t-t_n)/h_n)\|_{[W]_0(h_nI+t_n)} \\ &\quad\sim h_n^{1/2}\|U_n\|_{[W]_0(I)} \lesssim h_n^{1/2}\|U_n\|_{[W]^\bullet_2(I)} + \|P_{<1}v_n\|_{[W]_0(I)} + h_n^{1/2}\|P_{<h_n}(U_n - V_n)\|_{[W]_0(I)},\end{aligned} \tag{4-28}$$

where $\vec{V}_n := e^{it\langle\nabla\rangle_n}\vec{U}_n(0)$ and $\vec{v}_n = T_n\vec{V}_n(t/h_n)$. The first term on the right is vanishing since $\|U_n\|_{[W]^\bullet_2(I)}$ is bounded as shown above. The second term is $O(\delta)$ by Lemma 4.1. The third term is bounded — using Sobolev, Hölder and the same estimate as in (4-25) — by

$$|I|^{W_1}h_n^{1/2+d(1/2-W_2)}\|U_n - V_n\|_{L^\infty_t L^2_x(I)} \lesssim (|I|h_n)^{3/2-1/(d+1)}(\|\vec{U}_\infty\|_{L^\infty_t L^2_x(I)} + \varepsilon)^{p+1} = o(1), \tag{4-29}$$

hence (4-28) is $O(\delta)$ for large n. This concludes the proof of the lemma for $d \ge 6$.

The case $d \le 5$ is the same, but the nonlinear estimate is much simpler. In (4-13), $[\widetilde{M}]^\bullet_{2p}$ is replaced with $[M]_0$, and by the standard Strichartz we have

$$\|\gamma_n\|_{[W]^\bullet_2\cap[M]_0} + \|\overline{\gamma}_n\|_{L^\infty_t L^2_x} \lesssim \|\overline{\gamma}_n(0)\|_{L^2} + \|f'(U_\infty+\gamma_n) - f'(U_\infty)\|_{[W^{*(1)}]^\bullet_2} + \|h_n^2U_n\|_{L^1_t L^2_x} \tag{4-30}$$

and

$$\begin{aligned}\|f'(U_\infty+\gamma_n) - f'(U_\infty)\|_{[W^{*(1)}]^\bullet_2} &\lesssim \|(U_\infty, \gamma_n)\|^p_{[W]^\bullet_2\cap[M]_0}\|\gamma_n\|_{[W]^\bullet_2\cap[M]_0} \\ &\lesssim (\eta + \|\gamma_n\|_{[W]^\bullet_2\cap[M]_0})^p\|\gamma_n\|_{[W]^\bullet_2\cap[M]_0}.\end{aligned} \tag{4-31}$$

Then, estimating $\|h_n^2 U_n\|_{L_t^1 L_x^2(0,S)}$ in the same way as for $d \geq 6$, we obtain (4-14) without the last term. Equation (4-28) is the same as above. □

Proof of Theorem 3.1. Let v_n, V_n, V_∞ be the free solutions defined by

$$\vec{V}_n = e^{it\langle\nabla\rangle_n}\psi, \quad \vec{V}_\infty = e^{it|\nabla|}\psi, \quad \vec{v}_n = T_n V_n((t-t_n)/h_n), \tag{4-32}$$

and

$$M := \|U_\infty\|_{[W]_2^\bullet(J)}. \tag{4-33}$$

First consider the case $\tau_\infty = \infty$. Let $0 < \varepsilon < 1$ and choose $S > 0$ so large that

$$\delta_0 := \|V_\infty\|_{([W]_2^\bullet \cap [M]_0)(S,\infty)} \leq \delta(\varepsilon, M), \tag{4-34}$$

where $\delta(\cdot\,,\cdot)$ is given by Lemma 4.2. Then Lemma 4.1 implies that

$$\|v_n\|_{([W]_2 \cap [M]_0)(h_n S + t_n, \infty)} \lesssim \delta_0 \tag{4-35}$$

for large n. If $\delta_0 \ll 1$, then the standard scattering argument for NLKG using the Strichartz norms implies that $u_{(n)}$ exists on $(h_n S + t_n, \infty)$, satisfying

$$\|\vec{u}_{(n)} - \vec{v}_n\|_{L_t^\infty L_x^2(h_n S + t_n, \infty)} + \|u_{(n)} - v_n\|_{([W]_2 \cap [M]_0)(h_n S + t_n, \infty)} \lesssim \delta_0^{2^\star - 1} \ll \delta_0 \tag{4-36}$$

and also, for NLW,

$$\|\vec{U}_\infty - \vec{V}_\infty\|_{L_t^\infty L_x^2(S,\infty)} + \|U_\infty - V_\infty\|_{([W]_2^\bullet \cap [M]_0)(S,\infty)} \lesssim \delta_0^{2^\star - 1} \ll \delta_0. \tag{4-37}$$

Thus we obtain

$$\|u_{(n)}\|_{([W]_2 \cap [M]_0)(h_n S + t_n, \infty)} \lesssim \|V_\infty\|_{([W]_2^\bullet \cap [M]_0)(S,\infty)} \sim \|U_\infty\|_{([W]_2^\bullet \cap [M]_0)(S,\infty)} \tag{4-38}$$

and, for large n,

$$\|\vec{U}_{(n)}(S) - \vec{V}_n(S)\|_{L_x^2} + \|\vec{V}_n(S) - \vec{V}_\infty(S)\|_{L_x^2} + \|\vec{V}_\infty(S) - \vec{U}_\infty(S)\|_{L_x^2} \ll \delta_0. \tag{4-39}$$

The next step is to go from S to the negative time direction. If J is bounded from below, then let $S' := \inf J$. Otherwise, choose $S' < S$ so that

$$\|U_\infty\|_{([W]_2^\bullet \cap [M]_0)(-\infty, S')} < \varepsilon. \tag{4-40}$$

Applying Lemma 4.2 to U_∞ and $U_{(n)}$ backward in time from $t = S$, we obtain

$$\|\vec{U}_{(n)} - \vec{U}_\infty\|_{L_t^\infty L_x^2(S',S)} + \|U_{(n)} - U_\infty\|_{([W]_2^\bullet \cap [M]_0)(S',S)} < \varepsilon \tag{4-41}$$

and $\|u_{(n)}\|_{[W]_0(h_n S' + t_n, h_n S + t_n)} \lesssim \delta_0$ for large n.

If J is unbounded from below, we have still to go from S' to $-\infty$. The standard argument for small data scattering of NLW for $t \to -\infty$ implies that

$$\|\operatorname{Re} |\nabla|^{-1} e^{it|\nabla|} \vec{U}_\infty(S')\|_{([W]_2^\bullet \cap [M]_0)(-\infty,0)} \sim \|U_\infty\|_{([W]_2^\bullet \cap [M]_0)(-\infty,S')} < \varepsilon. \tag{4-42}$$

Then Lemma 4.1 applied backward in t implies, for large n,

$$\| \operatorname{Re} \langle\nabla\rangle^{-1} e^{it\langle\nabla\rangle} T_n \vec{U}_\infty(S') \|_{([W]_2 \cap [M]_0)(-\infty,0)} \lesssim \varepsilon. \tag{4-43}$$

Let w_n be the solution of NLKG with $\vec{w}_n(0) = T_n \vec{U}_{(n)}(S')$. Then the above estimate together with $\|\vec{U}_{(n)}(S') - \vec{U}_\infty(S')\|_{L^2_x} < \varepsilon$ and the scattering for NLKG implies

$$\|w_n\|_{([W]_2 \cap [M]_0)(-\infty,0)} \lesssim \varepsilon. \tag{4-44}$$

Since $w_n = h_n T_n U_{(n)}(t/h_n + S') = u_{(n)}(t + h_n S' + t_n)$, we deduce that

$$\begin{aligned} \|U_{(n)}\|_{([W]_2^\bullet \cap [M]_0)(-\infty,S')} &\sim \|u_{(n)}\|_{([W]_2^\bullet \cap [M]_0)(-\infty,h_n S'+t_n)} \\ &\lesssim \|u_{(n)}\|_{([W]_2 \cap [M]_0)(-\infty,h_n S'+t_n)} = \|w_n\|_{([W]_2 \cap [M]_0)(-\infty,0)} \lesssim \varepsilon. \end{aligned} \tag{4-45}$$

Thus we obtain, in the case $\tau_\infty = \infty$,

$$\|U_{(n)} - U_\infty\|_{([W]_2^\bullet \cap [M]_0)(J)} + \|u_n\|_{[W]_0(h_n J + t_n)} \lesssim \varepsilon + \delta_0 \tag{4-46}$$

for large n. Since ε and δ_0 can be chosen as small as we wish, this implies

$$\lim_{n\to\infty} \|U_{(n)} - U_\infty\|_{([W]_2^\bullet \cap [M]_0)(J)} + \|u_n\|_{[W]_0(h_n J + t_n)} = 0 \tag{4-47}$$

and, by scaling,

$$\|u_{(n)}\|_{([W]_2 \cap [M]_0)(h_n J + t_n)} \sim \|U_\infty\|_{([W]_2^\bullet \cap [M]_0)(J)} + \|u_{(n)}\|_{[W]_0(h_n J + t_n)} = \|U_\infty\|_{([W]_2^\bullet \cap [M]_0)(J)} + o(1). \tag{4-48}$$

Since $S \to \infty$ and $S' \to \inf J$ as $\varepsilon, \delta \to +0$, we also obtain

$$\lim_{n\to\infty} \|\vec{U}_{(n)} - \vec{U}_\infty\|_{L^\infty_t L^2_x(I)} = 0 \tag{4-49}$$

for any finite subinterval I. The case $\tau_\infty = -\infty$ is the same by the time symmetry.

If $\tau_\infty \in \mathbb{R}$ then $\|\vec{U}_{(n)}(\tau_\infty) - \vec{U}_\infty(\tau_\infty)\|_{L^2_x} \to 0$. Hence the same argument as we used above to go from S to $-\infty$ yields

$$0 = \lim_{n\to\infty} \|\vec{U}_{(n)} - \vec{U}_\infty\|_{L^\infty_t L^2_x(S',\tau_\infty)} = \lim_{n\to\infty} \|U_{(n)} - U_\infty\|_{([W]_2^\bullet \cap [M]_0)(\inf J,\tau_\infty)} \tag{4-50}$$

for any $S' \in (\inf J, \tau_\infty)$, and also on $(\tau_\infty, \sup J)$ by the time symmetry. Thus we obtain (4-47) and (4-49) for any $\tau_\infty \in [-\infty, \infty]$. □

Acknowledgments

The authors are grateful to Takahisa Inui, Tristan Roy, and Federica Sani for pointing out the errors.

References

[Ibrahim et al. 2011] S. Ibrahim, N. Masmoudi, and K. Nakanishi, "Scattering threshold for the focusing nonlinear Klein–Gordon equation", *Anal. PDE* **4**:3 (2011), 405–460. MR 2872122 Zbl 1270.35132

[Ibrahim et al. 2014] S. Ibrahim, N. Masmoudi, and K. Nakanishi, "Threshold solutions in the case of mass-shift for the critical Klein–Gordon equation", *Trans. Amer. Math. Soc.* **366**:11 (2014), 5653–5669. MR 3256178 Zbl 1302.35260

[Ibrahim et al. 2015] S. Ibrahim, N. Masmoudi, and K. Nakanishi, "Trudinger–Moser inequality on the whole plane with the exact growth condition", *J. Eur. Math. Soc.* **17**:4 (2015), 819–835. MR 3336837 Zbl 1317.35032

Received 20 Jun 2015. Accepted 6 Dec 2015.

SLIM IBRAHIM: ibrahim@math.uvic.ca
Department of Mathematics and Statistics, University of Victoria, PO Box 3060 STN CSC, Victoria, BC V8P 5C3, Canada

NADER MASMOUDI: masmoudi@courant.nyu.edu
Courant Institute for Mathematical Sciences, New York University, New York, NY 10012-1185, United States

KENJI NAKANISHI: nakanishi@ist.osaka-u.ac.jp
Department of Pure and Applied Mathematics, Osaka University, Graduate School of Information Science and Technology, Osaka 560-0043, Japan

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 9 No. 2 2016

2157-5045(2016)9:2;1-A