ANALYSIS & PDE

Volume 9 No. 1 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. 1, 2016
dx.doi.org/10.2140/apde.2016.9.1

RADEMACHER FUNCTIONS IN NAKANO SPACES

SERGEY ASTASHKIN AND MIECZYSŁAW MASTYŁO

The closed span of Rademacher functions is investigated in Nakano spaces $L^{p(\cdot)}$ on [0, 1] equipped with the Lebesgue measure. The main result of this paper states that under some conditions on distribution of the exponent function p the Rademacher functions form in $L^{p(\cdot)}$ a basic sequence equivalent to the unit vector basis in ℓ_2.

1. Introduction

We recall that the Rademacher functions on [0, 1] are defined by $r_k(t) = \operatorname{sign}(\sin 2^k \pi t)$ for every $t \in [0, 1]$ and each $k \in \mathbb{N}$. It is well known that (r_k) is an incomplete orthogonal system of independent random variables. This system plays a prominent role in the modern theory of Banach spaces and operators (see, e.g., [Diestel et al. 1995; Pisier 1986]). Special emphasis in this connection is placed on the study of local theory of Banach spaces and especially on using the notions of (Rademacher) type and cotype, which reflect the interplay between geometry and probability in these spaces. We mention here only a special case of the famous result due to Maurey and Pisier [1976]; it states that a Banach space has type strictly bigger than 1 (resp., finite cotype) if and only if it does not contain ℓ_1^n's (resp., ℓ_∞^n's) uniformly. For more details and a precise quantitative version of this result we refer, for example, to [Diestel et al. 1995, Chapter 14].

Rademacher functions play a significant role in the study of lattice and rearrangement-invariant structures in arbitrary Banach spaces. This research was initiated in the memoir [Johnson et al. 1979] by Johnson, Maurey, Schechtman and Tzafriri. By way of motivation let us also mention a classical result of Rodin and Semenov [1975], which states that the sequence (r_k) is equivalent in a symmetric space X to the unit vector basis in ℓ_2, that is,

$$\Big\| \sum_{k=1}^{\infty} a_k r_k \Big\|_X \approx \Big(\sum_{k=1}^{\infty} |a_k|^2 \Big)^{1/2}, \quad (a_k) \in \ell_2,$$

if and only if $G \subset X$, where G is the closure of $L^\infty[0, 1]$ in the Orlicz space $L_N[0, 1]$ generated by the function $N(t) = \exp(t^2) - 1$ for all $t \geq 0$. When this condition is satisfied, the span $[r_k]$ of Rademacher functions is complemented in X if and only if $X \subset G'$, where the Köthe dual space G' to G coincides (with equivalence of norms) with the Orlicz space $L_{N_*}[0, 1]$ generated by the Young conjugate N_* which

Astashkin's research was supported by the Ministry of Education and Science of Russian Federation. Mastyło's research was supported by the Foundation for Polish Science (FNP)..
MSC2010: primary 46E30; secondary 46B20, 46B42.
Keywords: Rademacher functions, Nakano spaces, symmetric spaces.

is equivalent at infinity to the function $t \mapsto t \log^{1/2} t$. This was proved independently by Rodin and Semenov [1979] and Lindenstrauss and Tzafriri [1979, pp. 134–138].

It is well known that (r_k) is a symmetric basic sequence in every symmetric space on [0, 1], however this is not true in the case of nonsymmetric Banach function lattices. In particular, this phenomenon takes place, for example, in the space of functions of bounded mean oscillation and as well as in Cesàro function spaces (see [Astashkin et al. 2011; Astashkin and Maligranda 2010]); this motivates searching for conditions under which Rademacher functions form a symmetric or an unconditional basic sequence in Banach function lattices.

The main purpose of this paper is to investigate the behaviour of Rademacher functions in the Nakano function spaces $L^{p(\cdot)}$ on [0, 1]. These spaces (which are also called "variable exponent Lebesgue spaces" in certain parts of the literature) are generalisations of the classical L^p-spaces, where the exponent p is allowed to vary measurably over a set of values in $[1, \infty)$.

Nakano spaces belong to the large family of Musielak–Orlicz spaces, and therefore many their basic properties follow from general results (see [Musielak 1983]). There are several books related to Nakano spaces, which cover some joint material, however, from somewhat different viewpoints. Let us mention [Diening et al. 2011] and [Cruz-Uribe and Fiorenza 2013], in which the authors provide a presentation of fundamentals of Nakano spaces and study whether certain principal results in modern harmonic analysis have natural analogues in the Nakano space setting. In the last decades the investigation on this topic has been also motivated by the modelling the so-called electrorheological fluids and some other applications (see [Cruz-Uribe and Fiorenza 2013], and also the more recent [Cruz-Uribe et al. 2014], where interesting connections between theory of Nakano spaces and strongly hyperbolic systems with time-dependent coefficients were discovered).

It is worth noting that a number of results related to the spaces $L^{p(\cdot)}$ is proved under some smoothness conditions on the exponent function p. Let us recall, as an example, a result of Sharapudinov [1986] which states that the Haar system is a basis in a Nakano space $L^{p(\cdot)}$ provided the exponent function p satisfies the piecewise Dini–Lipschitz condition with exponent $\alpha \geq 1$ (see also the above-cited [Diening et al. 2011; Cruz-Uribe and Fiorenza 2013]). In contrast to that in this paper we impose conditions upon distribution of p and investigate the problem whether they are sufficient or necessary for equivalence of the Rademacher sequence (r_k) in $L^{p(\cdot)}$ to the unit vector basis in ℓ_2.

2. Preliminaries

If (Ω, Σ, μ) is a σ-finite measure space, then, as usual, $L^0 := L^0(\mu)$ denotes the space of all real-valued μ-measurable functions. We say that $(X, \|\cdot\|_X)$ is a *Banach function lattice* (in short, *Banach lattice*) on (Ω, Σ, μ) if X is an ideal in L^0 and $\|f\|_X \leq \|g\|_X$ whenever $f, g \in X$ and $|f| \leq |g|$. The Köthe dual space X' of X is a collection of all elements $g \in L^0$ such that

$$\|g\|_{X'} := \sup\left\{ \int_\Omega |fg|\, d\mu;\ \|f\|_X \leq 1 \right\} < \infty.$$

The space $(X', \|\cdot\|_{X'})$ is a Banach function lattice with the Fatou property. Recall that a Banach function lattice X is said *to have the Fatou property* if the conditions $\sup_{n\geq 1}\|x_n\|_X < \infty$ and $x_n \to x$ a.e. imply that $x \in X$ and $\|x\|_X \leq \liminf_{n\to\infty} ||x_n||_X$. It is well known that X has the Fatou property if and only if the natural embedding of X into its second Köthe dual X'' is an isometric surjection.

Let $f \in L^0(I,m)$, where $I := [0,1]$ is equipped with the Lebesgue measure m. The *distribution function* of f is defined by $d_f(\lambda) = \mu(\{t \in I\,;\,|f(t)| > \lambda\})$, $\lambda \geq 0$, and its *decreasing rearrangement* by $f^*(t) = \inf\{s > 0; d_f(s) \leq t\}$, $t > 0$. One says that functions f and g are *equimeasurable* if $f^*(t) = g^*(t)$, $0 < t \leq 1$, or equivalently, $d_f(\lambda) = d_g(\lambda)$, $\lambda > 0$.

Recall some definitions and auxiliary results from the theory of symmetric spaces (for more details see [Bennett and Sharpley 1988; Kreĭn et al. 1982]).

A Banach function lattice X on (I,m) is called a *symmetric space* if the conditions $f^* \leq g^*$ a.e. on I and $g \in X$ imply $f \in X$ and $\|f\|_X \leq \|g\|_X$. The fundamental function of a symmetric space X is given by $\varphi_X(t) = \|\chi_{[0,t)}\|_X$ for all $t \in I$. In what follows we will use the following obvious inequality for any symmetric space X on I,

$$f^*(t) \leq \frac{1}{\varphi_X(t)}\,\|f\|_X, \quad f \in X,\ t \in (0,1]. \tag{1}$$

Important examples of symmetric spaces are Orlicz, Marcinkiewicz and Lorentz spaces. Recall that $\Phi : [0,\infty) \to [0,\infty)$ is called an Orlicz function if $\Phi(0) = 0$ and Φ is positive, nondecreasing, convex and left-continuous on $(0,\infty)$. If Φ is such a function, the Orlicz space L_Φ consists of all $f \in L^0(m)$ for which there exists $\lambda > 0$ such that

$$\int_I \Phi\big(|f|/\lambda\big)\,dm < \infty.$$

It is a symmetric space equipped with the norm

$$\|f\|_{L_\Phi} = \inf\left\{\lambda > 0;\ \int_I \Phi\left(\frac{|f|}{\lambda}\right)dm \leq 1\right\}.$$

In what follows by L_N(resp., L_M) we will denote the Orlicz space on $[0,1]$ generated by the function $N(t) = \exp(t^2) - 1$ (resp., $M(t) = \exp(t^2\log(t+1)) - 1$) for all $t \geq 0$.

Let $\varphi : I \to [0,\infty)$ be a quasiconcave function, that is $\varphi(0) = 0$, $\varphi(t) > 0$ for $t \in I$ and both φ and $t \mapsto \widetilde{\varphi}(t) := t/\varphi(t)$ are nondecreasing functions on $(0,1]$. The Marcinkiewicz space $M(\varphi)$ is defined to be the space of all $f \in L^0(m)$ equipped with the norm

$$\|f\|_{M(\varphi)} = \sup_{0<s\in I}\frac{1}{\varphi(s)}\int_0^s f^*(t)\,dt.$$

If $\varphi : I \to [0,\infty)$ is an increasing concave function, $\varphi(0) = 0$, the Lorentz space $\Lambda(\varphi)$ consists of all $f \in L^0$ such that

$$\|f\|_{\Lambda(\varphi)} = \int_0^1 f^*(t)\,d\varphi(t) < \infty.$$

It is well known that L^1 and L^∞ are, respectively, the largest and the smallest symmetric spaces on I; moreover, if X is a symmetric space on I with the fundamental function φ, then φ is quasiconcave and the following continuous embeddings hold (see [Kreĭn et al. 1982, Theorems II.5.5 and II.5.7] or [Bennett and Sharpley 1988, Theorem II.5.13]):

$$\Lambda(\bar{\varphi}) \hookrightarrow X \hookrightarrow M(\widetilde{\varphi}),$$

where $\bar{\varphi}$ is the least concave majorant of φ. In what follows we will frequently use the well-known fact that the Orlicz space L_N generated by the function $N(t) = \exp(t^2) - 1$, $t \geq 0$, coincides up to equivalence of norms with the Marcinkiewicz space $M(\varphi)$ generated by the function $\varphi(t) = t\log^{1/2}(e/t)$, $0 < t \leq 1$ (see [Lorentz 1951]).

Let (Ω, Σ, μ) be a σ-finite measure space. Given a measurable function $p : \Omega \to [1, \infty)$, we define the Nakano space $L^{p(\cdot)}(\mu)$ to be the space of all $f \in L^0(\mu)$ such that for some $\lambda > 0$

$$\rho_\lambda(f) = \int_\Omega \left(\frac{|f(t)|}{\lambda}\right)^{p(t)} d\mu < \infty.$$

$L^{p(\cdot)}(\mu)$ becomes a Banach function lattice with the Fatou property when equipped with the norm

$$\|f\|_{L^{p(\cdot)}} = \|f\|_{p(\cdot)} := \inf\{\lambda > 0; \rho_\lambda(f/\lambda) \leq 1\}.$$

Throughout the paper a Nakano space defined on $[0, 1]$ equipped with the Lebesgue measure m is denoted for short $L^{p(\cdot)}$. Notice that $L^{p(\cdot)}$ is not a symmetric space unless the exponent p is a constant function, and in this case we write $\|\cdot\|_p$ instead of $\|\cdot\|_{L^p}$.

Further, we shall frequently use the following lemma which is an immediate consequence of Theorem 3 from [Fiorenza and Rakotoson 2007].

Lemma 2.1. *Let $f : [0, 1] \to [0, \infty)$ and $p : [0, 1] \to [1, \infty)$ be two Lebesgue measurable functions. Then*

$$\|f\|_{L^{p(\cdot)}} \leq 4\|f^*\|_{L^{p^*(\cdot)}}.$$

3. Main results

In this section we shall prove the main results of the paper. We recall that L_N and L_M are the Orlicz spaces on $[0, 1]$ generated by the functions $N(t) = \exp(t^2) - 1$ and $M(t) = \exp(t^2 \log(t + 1)) - 1$).

Theorem 3.1. *Let $p : (0, 1] \to [1, \infty)$ be a Lebesgue measurable function and let $L^{p(\cdot)}$ be the Nakano space generated by p. Each of the following conditions implies the next*:

(i) $L_N \subset L^{p(\cdot)}$.

(ii) *The Rademacher system (r_n) is equivalent in the space $L^{p(\cdot)}$ to the unit vector basis in ℓ_2.*

(iii) *There is a constant $C > 0$ such that*

$$m(\{t \in [0, 1];\ p(t) > \lambda\}) \leq C^\lambda \lambda^{-\lambda/2}, \quad \lambda \geq 1.$$

(iv) $L_M \subset L^{p(\cdot)}$.

We start with the following distribution estimate, which will be useful for us in the sequel:

Proposition 3.1. *Suppose that for each $k \in \mathbb{N}$ and $m \in \mathbb{N}$ there exists $\ell > m$ such that*

$$\Big\| \sum_{i=\ell+1}^{\ell+k} r_i \Big\|_{L^{p(\cdot)}} \le B\sqrt{k},$$

where $B > 0$ is independent of k and m. Then

$$m(\{t \in [0,1];\ p(t) > \lambda\}) \le 2(4B)^{\lambda}\lambda^{-\lambda/2}, \quad \lambda \ge 1.$$

Proof. Let $\lambda \ge 1$ be fixed. We put

$$E_\lambda := \{t \in [0,1]; p(t) > \lambda\}.$$

Without loss of generality, we can assume that $m(E_\lambda) > 0$. By the Sagher–Zhou local version of Khintchine inequality for L^1 (see [Sagher and Zhou 1990, Theorem 1]), it follows that there exists $n(\lambda)$ such that for all $n \ge n(\lambda)$, every Rademacher sum $R_n = \sum_{k=n}^{\infty} a_k r_k$ and arbitrary $(a_k) \in \ell_2$ with $\|(a_k)\|_{\ell_2} = 1$, we have

$$\int_{E_\lambda} |R_n(t)|\, dt \ge \alpha\, m(E_\lambda),$$

where $\alpha > 0$ is a universal constant. Since $\lambda \ge 1$,

$$\left(\frac{1}{m(E_\lambda)} \int_{E_\lambda} |R_n(t)|^{\lambda}\, dt\right)^{1/\lambda} \ge \frac{1}{m(E_\lambda)} \int_{E_\lambda} |R_n(t)|\, dt,$$

and so

$$\left(\int_{E_\lambda} |R_n(t)|^{\lambda}\, dt\right)^{1/\lambda} \ge \alpha (m(E_\lambda))^{1/\lambda}. \tag{2}$$

On the other hand, it is well known (in particular, it is a consequence of the above-cited Rodin–Semenov theorem) that there exists a constant $\beta > 0$ such that

$$\Big\| \sum_{k=1}^{\infty} a_k r_k \Big\|_{L_N} \le \beta\, \|(a_k)\|_{\ell_2}, \quad (a_k) \in \ell_2, \tag{3}$$

where, as above, L_N is the Orlicz space generated by the function $N(t) = \exp(t^2) - 1$, $t \ge 0$. Since the fundamental function of L_N is given by $\varphi(t) = 1/N^{-1}(1/t) = \log^{-1/2}(1 + 1/t)$ for all $t \in (0,1]$, it follows by (1) and (3) that

$$\left(\sum_{n=1}^{\infty} a_k r_k\right)^{*}(t) \le \beta \log^{1/2}\left(1 + \frac{1}{t}\right) \le \beta \log^{1/2}\left(\frac{e}{t}\right), \quad t \in (0,1],$$

for all $(a_k) \in \ell_2$ with $\|(a_k)\|_{\ell_2} \le 1$. Hence, for every $\delta > 0$ and $E \subset [0,1]$ with $m(E) < \delta$, we obtain

$$\left(\int_E |R_n(t)|^{\lambda}\, dt\right)^{1/\lambda} \le \left(\int_0^{\delta} R_n^{*}(t)^{\lambda}\, dt\right)^{1/\lambda} \le \beta \left(\int_0^{\delta} \log^{\lambda/2}\left(\frac{e}{t}\right) dt\right)^{1/\lambda}.$$

Choose $\delta = \delta(\lambda) > 0$ so that

$$\int_0^\delta \log^{\lambda/2}\Big(\frac{e}{t}\Big)\,dt \le \beta^{-\lambda}\alpha^\lambda m(E_\lambda).$$

Then, from the preceding inequality and (2), it follows that

$$\Big(\int_E |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} \le \alpha(m(E_\lambda))^{1/\lambda} \le \Big(\int_{E_\lambda} |R_n(t)|^\lambda\,dt\Big)^{1/\lambda}. \tag{4}$$

provided $m(E) < \delta$ and $\|(a_k)\|_{\ell_2} = 1$.

We denote by I_k^ν the dyadic interval $[(k-1)2^{-\nu}, k2^{-\nu}]$ for each $\nu \in \mathbb{Z}_+$ and each $1 \le k \le 2^\nu$. Then we can find a finite union of pairwise disjoint intervals $F = \bigcup_{j=1}^m I_{k_j}^{\nu_j}$, $1 \le k_j \le 2^{\nu_j}$, $1 \le j \le m$ such that

$$m(E_\lambda \triangle F) \le \max\big\{\delta, \tfrac12 m(E_\lambda)\big\}$$

(here, $A \triangle B := (A \setminus B) \cup (B \setminus A)$). Hence, $m(F) \ge m(E_\lambda) - m(E_\lambda \triangle F) \ge \frac12 m(E_\lambda)$, and for each sum $R_n = \sum_{k=n}^\infty a_k r_k$ with $\|(a_k)\|_{\ell_2} = 1$, by (4), we obtain

$$\Big(\int_F |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} \le \Big(\int_{E_\lambda} |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} + \Big(\int_{E_\lambda \triangle F} |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} \le 2\Big(\int_{E_\lambda} |R_n(t)|^\lambda\,dt\Big)^{1/\lambda}.$$

This implies that

$$\Big(\frac{1}{m(E_\lambda)}\int_{E_\lambda} |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} \ge \frac12\Big(\frac{1}{2m(F)}\int_F |R_n(t)|^\lambda\,dt\Big)^{1/\lambda} \ge \frac14\Big(\frac{1}{m(F)}\int_F |R_n(t)|^\lambda\,dt\Big)^{1/\lambda}.$$

Now, let a positive integer $m \ge n(\lambda)$ be such that all Rademacher functions r_k with $k \ge m$ change their sign at least once on each dyadic component of the set F. Then for any $(a_k) \in \ell_2$,

$$\Big(\frac{1}{m(F)}\int_F \Big|\sum_{k=m}^\infty a_k\, r_k(t)\Big|^\lambda dt\Big)^{1/\lambda} = \Big\|\sum_{k=m}^\infty a_k\, r_k\Big\|_\lambda.$$

Combining this equality with the above estimate, we obtain

$$\Big(\frac{1}{m(E_\lambda)}\int_{E_\lambda} |R_m(t)|^\lambda\,dt\Big)^{1/\lambda} \ge \tfrac14\|R_m\|_\lambda \tag{5}$$

for every sum $R_m = \sum_{k=m}^\infty a_k r_k$, $\|(a_k)\|_{\ell_2} = 1$ (m depends on λ). Our hypothesis implies that for each $\lambda \ge 1$ we can find $\ell > m$ such that

$$\Big\|\sum_{i=\ell+1}^{\ell+[\lambda]} r_i\Big\|_{L^{p(\cdot)}} \le B\sqrt{[\lambda]}, \tag{6}$$

where, as usual, $[x]$ is the integer part of x. In the opposite direction, we will use the following well-known inequality (see, e.g., [Blei 2001, Lemma VII.30, p. 167]):

$$2\left\|\sum_{j=1}^{k} r_j\right\|_k \geq k, \quad k \in \mathbb{N}.$$

If $R_{\lambda,\ell} := \sum_{i=\ell+1}^{\ell+[\lambda]} r_i$, this inequality yields

$$2\,\|R_{\lambda,\ell}\|_\lambda \geq 2\,\|R_{\lambda,\ell}\|_{[\lambda]} = 2\left\|\sum_{j=1}^{[\lambda]} r_j\right\|_{[\lambda]} \geq 2\,[\lambda].$$

Let $\bar{R}_{\lambda,\ell} := R_{\lambda,\ell}/\sqrt{[\lambda]}$. Then, from the latter inequality it follows that

$$\|\bar{R}_{\lambda,\ell}\|_{[\lambda]} \geq 2\sqrt{[\lambda]} \geq \sqrt{\lambda}.$$

Moreover, it is easy to see that $\bar{R}_{\lambda,\ell} = \sum_{k=m}^{\infty} a'_k r_k$, with $\|(a'_k)\|_{\ell_2} = 1$. Combining the preceding estimate with inequality (5), we obtain

$$\left(\frac{1}{m(E_\lambda)}\int_{E_\lambda} \left|\bar{R}_{\lambda,\ell}(t)\right|^\lambda dt\right)^{1/\lambda} \geq \tfrac{1}{4}\sqrt{\lambda},$$

or equivalently,

$$\|\bar{R}_{\lambda,\ell}\,\chi_{E_\lambda}\|_\lambda \geq \tfrac{1}{4}\sqrt{\lambda}\, m(E_\lambda)^{1/\lambda}. \tag{7}$$

where χ_{E_λ} is the characteristic function of the set E_λ. On the other hand, in view of (6) we have $\|\bar{R}_{\lambda,\ell}\|_{L^{p(\cdot)}} \leq B$ and so, setting $\bar{E}_\lambda = \{t \in E_\lambda;\ |\bar{R}_{\lambda,\ell}(t)| \geq B\}$, by the definition of the norm in the Nakano space $L^{p(\cdot)}$, we deduce

$$\int_{E_\lambda}\left|\frac{\bar{R}_{\lambda,\ell}(t)}{B}\right|^\lambda dt \leq \int_{\bar{E}_\lambda}\left|\frac{\bar{R}_{\lambda,\ell}(t)}{B}\right|^\lambda dt + \int_{E_\lambda\setminus\bar{E}_\lambda}\left|\frac{\bar{R}_{\lambda,\ell}(t)}{B}\right|^\lambda dt \leq \int_0^1\left|\frac{\bar{R}_{\lambda,\ell}(t)}{B}\right|^{p(t)} dt + 1 \leq 2. \tag{8}$$

Therefore, from (7) it follows that

$$2 \geq \int_{E_\lambda}\left|\frac{\bar{R}_{\lambda,\ell}(t)}{B}\right|^\lambda dt \geq \frac{\lambda^{\lambda/2}}{(4B)^\lambda} m(E_\lambda),$$

whence $m(E_\lambda) \leq 2(4B)^\lambda \lambda^{-\lambda/2}$. This completes the proof. □

Proof of Theorem 3.1. (i) ⇒ (ii). First, by [Diening et al. 2011, Theorem 3.3.1], for any exponent $p(\cdot)$ we have

$$\|f\|_{L^1} \leq 2\,\|f\|_{L^{p(\cdot)}}, \quad f \in L^{p(\cdot)}.$$

Combining this with the Khintchine inequality in L^1 (see [Szarek 1976]), we obtain

$$\left\|\sum_{k=1}^{\infty} a_k\, r_k\right\|_{L^{p(\cdot)}} \geq \frac{1}{2\sqrt{2}}\,\|(a_k)\|_{\ell_2}, \quad (a_k) \in \ell_2.$$

Thus our hypothesis and (3) imply that there exists a constant $C > 0$ such that

$$\Big\| \sum_{k=1}^{\infty} a_k\, r_k \Big\|_{L^{p(\cdot)}} \le C\, \|(a_k)\|_{\ell_2}, \quad (a_k) \in \ell_2.$$

The implication (ii) ⇒ (iii) follows from Proposition 3.1.

(iii) ⇒ (iv). Note that the Orlicz space L_M, where $M(t) = \exp(t^2 \log(1+t)) - 1$ for all $t \ge 0$, coincides with the Marcinkiewicz space with the fundamental function $\varphi := \varphi_{L_M}$ given by

$$\varphi(t) := \left(\frac{\log(e/t)}{\log\log(e^2/t)} \right)^{-1/2}, \quad 0 < t \le 1$$

(see, e.g., [Lorentz 1951] or [Astashkin 2009, Lemma 3.2]). Hence, L_M can be characterised as the set of all measurable functions x on $[0, 1]$ for which there exists a constant $C > 0$ such that

$$x^*(t) \le \frac{C}{\varphi(t)}, \quad 0 < t \le 1.$$

Thus, since $L^{p(\cdot)}$ is a Banach lattice, the embedding $L^{p(\cdot)} \supset L_M$ will be proved if we show that the space $L^{p(\cdot)}$ contains all functions equimeasurable with the function

$$f_0(t) = \frac{1}{\varphi(t)}, \quad 0 < t \le 1.$$

By hypothesis and Lemma 2.1, it follows that we need only to check that for some $\lambda > 0$

$$\int_0^1 \left(\frac{f_0(t)}{\lambda} \right)^{g(t)} dt < \infty, \tag{9}$$

where g is a decreasing positive function on $(0, 1]$ such that $g(t) \ge 1$ and

$$m(\{t \in (0,1];\ g(t) > x\}) = g^{-1}(x) = C^x x^{-x/2}, \quad x \ge 1,$$

for some $C \ge 1$.

For $x_0 \ge 1$, which can be chosen later, we have

$$\begin{aligned}\int_0^{g^{-1}(x_0)} \left(\frac{f_0(t)}{\lambda} \right)^{g(t)} dt &= -\int_{x_0}^{\infty} \left(\frac{f_0(C^x x^{-x/2})}{\lambda} \right)^{x} d(C^x x^{-x/2}) \\ &= \int_{x_0}^{\infty} \left(\frac{f_0(C^x x^{-x/2})}{\lambda} \right)^{x} C^x x^{-x/2} \log(C^{-1} e^{1/2} x^{1/2})\, dx.\end{aligned}$$

If x_0 is sufficiently large, then for all $x \ge x_0$ we infer

$$f_0(C^x x^{-x/2}) = \left(\frac{\log(e C^{-x} x^{x/2})}{\log\log(e^2 C^{-x} x^{x/2})} \right)^{1/2} = \frac{1}{\sqrt{2}} \left(\frac{x \log(C^{-x} e^{2/x} x)}{\log x + \log\left(\frac{1}{2}\log(C^{-2} e^{4/x} x)\right)} \right)^{1/2} \le x^{1/2}.$$

Therefore, the preceding inequality implies

$$\int_0^{g^{-1}(x_0)} \left(\frac{f_0(t)}{\lambda}\right)^{g(t)} dt \le \int_{x_0}^{\infty} \left(\frac{C}{\lambda}\right)^{x} \log(C^{-1}e^{1/2}x^{1/2})\, dx < \infty,$$

provided that $\lambda > C$. Clearly, we obtain (9).

Finally, implication (iv) $\Rightarrow$ (i) is an immediate consequence of the obvious embedding $L_N \subset L_M$, and the proof is complete. □

We do not know whether the distribution condition from (iii) implies the embedding $L_N \subset L^{p(\cdot)}$ or the equivalence of Rademacher system in $L^{p(\cdot)}$ to the unit vector basis in ℓ_2. However, the next result can be treated as an approach to the solution of these problems. In its first part we prove that some stronger condition on the distribution function of an exponent $p(\cdot)$ insures the embedding $L_N \subset L^{p(\cdot)}$ and in the second one we show that this result is in a sense sharp.

Theorem 3.2. *Let $p : (0, 1] \to [1, \infty)$ be a Lebesgue measurable function.*

(a) *If there exists a constant $C > 0$ such that*

$$m(\{t \in (0,1];\, p(t) > x\}) \le C^x (x \log x)^{-x/2}, \quad x \ge 1,$$

then $L_N \subset L^{p(\cdot)}$.

(b) *If there exists an increasing differentiable function θ such that $\lim_{x\to\infty} \theta(x) = \infty$, the function $x \mapsto \theta(x)x^{-1/2}\log^{-1/2} x$ is decreasing for large enough x, and $\liminf_{x\to\infty} m(\{t \in (0,1];\ p(t) > x\}\, \theta(x)^{-x} x^{x/2} \log^{x/2} x > 0$,*

then $L_N \not\subset L^{p(\cdot)}$.

Proof. (a) It can be easily checked that the function $x \mapsto C^x(x\log x)^{-x/2}$ decreases if $x \ge x_0$, where $x_0 > 1$ is sufficiently large. Denote by q the function inverse to it on the interval $[0, t_0]$, where $q(t_0) = x_0$. Then, from our hypothesis on p, it follows that $p^*(t) \le q(t)$ for all $0 < t \le t_0$. Recall that the space L_N coincides with the Marcinkiewicz space whose fundamental function is given by $t \mapsto \log^{-1/2}(e/t)$, $t \in (0, 1)$. Therefore, thanks to Lemma 2.1, we need only to check that for some $\lambda > 0$

$$I_\lambda := \int_0^{t_0} \left(\frac{\log^{1/2}(e/t)}{\lambda}\right)^{q(t)} dt < \infty.$$

In fact,

$$\begin{aligned} I_\lambda &= -\int_{x_0}^{\infty} (\lambda^{-1}\log^{1/2}(eC^{-x}(x\log x)^{x/2}))^x \, d(C^x(x\log x)^{-x/2}) \\ &= \frac{1}{2}\int_{x_0}^{\infty} \lambda^{-x}\left(\frac{x}{2}\right)^{x/2} \log^{x/2}(e^{2/x}C^{-2}x\log x)\cdot C^x(x\log x)^{-x/2}\left(\log(C^{-2}x\log x) + \frac{\log x + 1}{\log x}\right) dx \\ &\le C_1 \int_{x_0}^{\infty} \left(\frac{C}{\lambda}\right)^x \left(\log(x\log x) + \frac{\log x + 1}{\log x}\right) dx < \infty, \end{aligned}$$

provided $\lambda > C$, and this completes the proof.

(b) It is sufficient to show that for every $\lambda > 0$ there exists a measure-preserving transformation ω of $(0, 1]$ such that

$$\int_0^1 \left(\lambda^{-1} \log^{1/2}(e/\omega(t))\right)^{p(t)} dt = \infty. \tag{10}$$

In fact, from (10) it follows that $\log^{1/2}(e/\omega) \notin L^{p(\cdot)}$. On the other hand, since ω preserves measure, we have

$$\left(\log^{1/2}\left(\frac{e}{\omega(\cdot)}\right)\right)^*(t) = \log^{1/2}\left(\frac{e}{t}\right), \quad t \in (0, 1].$$

Combining this with the fact that $L_N = M(\varphi)$, where $\varphi(t) = t\log^{1/2}(e/t)$, $0 < t \le 1$, we infer $\log^{1/2}(e/\omega) \in L_N$ and the desired result follows.

Let us prove (10). Without loss of generality, we can assume that

$$\theta(x) \le \log^{1/2} x, \text{ for large enough } x \tag{11}$$

(otherwise, instead of $\theta(x)$ we can take the function $\min\{\theta(x), \log^{1/2} x\}$). Moreover, our hypotheses on θ imply

$$\left(\frac{\theta(x)^2}{x \log x}\right)' = x^{-2} \log^{-2} x (2\theta'(x)\theta(x)x \log x - \theta^2(x)(1 + \log x)) \le 0,$$

and so

$$\frac{2x\theta'(x)}{\theta(x)} \le \frac{1 + \log x}{x}, \quad x \ge x_0, \tag{12}$$

if $x_0 \ge 1$ is sufficiently large.

By assumption, there exists $\alpha \in (0, 1)$ such that for all $x \ge x_0$ we have

$$m\{t \in (0, 1];\ p(t) > x\} \ge \alpha \psi(x)^x.$$

Hence, if g is the inverse function to the mapping $x \mapsto \alpha\psi(x)^x$, $x \ge x_0$, we obtain

$$p^*(t) \ge g(t), \quad 0 < t \le t_0, \tag{13}$$

for some $t_0 \in (0, 1]$. If it is necessary, diminishing t_0 we can assume also, for a given $\lambda > 0$, the inequality $\log^{1/2}(e/t) \ge \lambda$ to be valid for all $t \in (0, t_0]$.

Let ω be a measure-preserving transformation of $(0, 1]$ such that $p(t) = p^*(\omega(t))$ (see [Bennett and Sharpley 1988, Theorem 2.7.5]). From inequality (13) it follows that

$$p(t) \ge g(\omega(t)), \quad t \in E,$$

where $E = \omega^{-1}([0, t_0])$. As a consequence,

$$\begin{aligned} I_\lambda &:= \int_E \left(\lambda^{-1} \log^{1/2}(e/\omega(t))\right)^{p(t)} dt \ge \int_E \left(\lambda^{-1} \log^{1/2}(e/\omega(t))\right)^{g(\omega(t))} dt \\ &= \int_0^{t_0} (\lambda^{-1} \log^{1/2}(e/t))^{g(t)} dt, \end{aligned}$$

and by letting $x = g(t)$, we obtain

$$I_\lambda \ge -\alpha \int_{g(t_0)}^{\infty} \lambda^{-x} \log^{x/2}\Big(\frac{e}{\alpha\psi(x)^x}\Big)\, d(\psi(x)^x).$$

Together with the elementary calculations

$$\begin{aligned}(\psi(x)^x)' &= \Big(\exp\Big(-\frac{x}{2}\log(\theta(x)^{-2}x\log x)\Big)\Big)' \\ &= \psi(x)^x\Big(-\frac{1}{2}\log(\theta(x)^{-2}x\log x)-\frac{x}{2}\,\frac{\theta(x)^2}{x\log x}\theta^{-4}(x)\big((1+\log x)\theta^2(x)-2\theta(x)\theta'(x)x\log x\big)\Big) \\ &= -\frac{1}{2}\psi(x)^x\Big(\log\frac{x\log x}{\theta^2(x)}+\frac{1+\log x}{\log x}-\frac{2x\theta'(x)}{\theta(x)}\Big),\end{aligned}$$

inequality (12) shows that

$$(\psi(x)^x)' \le -\frac{1}{2}\psi(x)^x \log\frac{x\log x}{\theta^2(x)}, \quad x \ge x_0.$$

Combining this with the preceding inequality and (11), we obtain

$$\begin{aligned}I_\lambda &\ge \frac{\alpha}{2}\int_{g(t_0)}^{\infty} \lambda^{-x}\Big(\frac{x}{2}\Big)^{x/2} \log^{x/2}(\alpha^{-2/x}e^{2/x}\theta(x)^{-2}x\log x)\,\theta(x)^x x^{-x/2}\log^{-x/2} x\ \log\frac{x\log x}{\theta^2(x)}\,dx \\ &\ge \frac{\alpha}{2}\int_{g(t_0)}^{\infty}(\lambda\sqrt{2})^{-x}\theta(x)^x\log x\,dx.\end{aligned}$$

Since $\lim_{x\to\infty}\theta(x) = \infty$, from the last estimate it follows that $I_\lambda = \infty$, which implies (10).

The proof is complete. □

We conclude the paper with the result which can be treated as a complement to Theorem 3.1 showing that equivalence of the Rademacher system in $L^{q(\cdot)}$ with arbitrary exponent q, which is equimeasurable with a given p, to the unit vector basis in ℓ_2 implies the embedding $L_N \subset L^{p(\cdot)}$.

Given a Lebesgue measurable function $p : [0, 1] \to [1, \infty)$ we let $\Omega(p)$ to be the set of all functions $q \in L^0(m)$ which are equimeasurable with p.

Theorem 3.3. *Suppose that for every $q \in \Omega(p)$ the Rademacher system is equivalent in the space $L^{q(\cdot)}$ to the standard basis in ℓ_2. Then $L_N \subset L^{q(\cdot)}$ for every $q \in \Omega(p)$.*

Proof. Our hypothesis yields that for any $q \in \Omega(p)$ there exits a constant $C_q > 0$ such that for every $a = (a_k) \in \ell_2$

$$\Big\|\sum_{k=1}^{\infty} a_k r_k\Big\|_{L^{q(\cdot)}} \le C_q\,\|a\|_{\ell_2}. \tag{14}$$

We claim that there is a constant $C_0 > 0$ such that for every measure-preserving mapping $\omega : [0, 1] \to [0, 1]$

and all $a = (a_k) \in \ell_2$ we have

$$\left\| \sum_{k=1}^{\infty} a_k r_k \right\|_{L^{p^*(\omega(\cdot))}} \le C_0 \, \|a\|_{\ell_2}. \tag{15}$$

To see this we define the linear operator $T_\omega : \ell_2 \to L^{p^*(\cdot)}$:

$$T_\omega(a_k) := \sum_{k=1}^{\infty} a_k \, r_k(\omega^{-1}), \qquad (a_k) \in \ell_2,$$

generated by an arbitrary measure-preserving mapping $\omega : [0,1] \to [0,1]$. Since for any $\lambda > 0$

$$\int_0^1 \left| \frac{1}{\lambda} \, T_\omega a(t) \right|^{p^*(t)} dt = \int_0^1 \left(\frac{1}{\lambda} \left| \sum_{k=1}^{\infty} a_k r_k(t) \right| \right)^{p^*(\omega(t))} dt \tag{16}$$

and the function $q := p^*(\omega) \in \Omega(p)$, from (14) it follows that the operator T_ω is bounded from ℓ_2 into $L^{p^*(\cdot)}$.

For a given sequence $b = (b_k) \in \ell_2$ we let $f = \left| \sum_{k=1}^{\infty} b_k \, r_k \right|$. Applying Theorem 2.7.5 from [Bennett and Sharpley 1988] once more, we can find a measure-preserving mapping $v : [0,1] \to [0,1]$ such that $f = f^*(v)$. Since $p^*(v) \in \Omega(p)$, by (14), we have

$$\|f\|_{L^{p^*(v)}} \le K := C_{p^*(v)} \, \|b\|_{\ell_2}.$$

Therefore,

$$\int_0^1 \left(\frac{f^*(t)}{K} \right)^{p^*(t)} dt = \int_0^1 \left(\frac{f(v^{-1}(t))}{K} \right)^{p^*(t)} dt = \int_0^1 \left(\frac{f(t)}{K} \right)^{p^*(v(t))} dt \le 1,$$

whence, by Lemma 2.1,

$$\int_0^1 \left(\frac{f(t)}{K} \right)^{p^*(\omega(t))} dt = \int_0^1 \left(\frac{f(\omega^{-1}(t))}{K} \right)^{p^*(t)} dt \le 3.$$

Combining the last inequality and equality (16), with $a = b$, we get

$$\|T_\omega b\|_{L^{p^*(\cdot)}} \le 3K = 3 \, C_{p^*(v)} \|b\|_{\ell_2},$$

where the constant $C_{p^*(v)}$ does not depend on ω. Thus, the family of operators $\{T_\omega\}_{\omega \in \Omega(p)}$ is pointwise bounded, and thanks to the uniform boundedness principle, we obtain

$$\|T_\omega a\|_{L^{p^*(\cdot)}} \le C_0 \|a\|_{\ell_2}$$

for some constant C_0 independent of ω. Clearly, inequality (15) is an immediate consequence of the latter inequality and (16).

Let us continue the proof of Theorem 3.3. As above, G is the closure L^∞ in the Orlicz space L_N. By [Astashkin and Semënov 2013, Theorem 4], for arbitrary $x \in G$ there exists a Rademacher sum

$f_1 = \sum_{k=1}^{\infty} a_k r_k$ such that

$$\|a\|_{\ell_2} \le C_1 \|x\|_{L_N} \quad \text{and} \quad x^*(t) \le C_2(\|a\|_{\ell_2} + f_1^*(t)), \qquad t \in (0, 1]. \tag{17}$$

Take a measure-preserving mapping $\omega : [0, 1] \to [0, 1]$, for which $|f_1| = f_1^*(\omega)$. Then, from (17) and (15) it follows

$$\begin{aligned}\|x^*\|_{L^{p^*(\cdot)}} &\le C_2(\|a\|_{\ell_2} + \|f_1(\omega^{-1})\|_{L^{p^*(\cdot)}}) = C_2(\|a\|_{\ell_2} + \|f_1\|_{L^{p^*(\omega)}}) \\ &\le C_2(1 + C_0)\,\|a\|_{\ell_2} \le C_1 C_2(1 + C_0)\|x\|_{L_N}.\end{aligned}$$

Furthermore, letting $x_n(t) = \min\{n, \log^{1/2}(e/t)\}$, $t \in (0, 1]$, we have $x_n = x_n^* \in G$ and $\|x_n\|_{L_N} \le \alpha := \|\log^{1/2}(e/t)\|_{L_N}$ for each $n \in \mathbb{N}$. Hence, from the previous inequality it follows that

$$\|x_n\|_{L^{p^*(\cdot)}} \le C_1 C_2(1 + C_0)\alpha, \quad n \in \mathbb{N}.$$

Since the space $L^{p^*(\cdot)}$ has the Fatou property and $\lim_{t\to\infty} x_n(t) = \log^{1/2}(e/t)$, we infer that the function $t \mapsto \log^{1/2}(e/t)$ lies in $L^{p^*(\cdot)}$. Recall that L_N consists of all $x \in L^0(m)$ such that $x^*(t) \le C \log^{1/2}(e/t)$ for all $t \in (0, 1]$ and some constant $C > 0$. Therefore, by Lemma 2.1, we obtain $L_N \subset L^{p^*(\cdot)}$. Combining this with the fact that L_N is a symmetric space, we deduce $L_N \subset L^{q(\cdot)}$ for arbitrary exponent $q \in \Omega(p)$, which completes the proof. ☐

Let us observe that, if a function p satisfies the conditions of Theorem 3.2(b), the Rademacher system (r_n) in $L^{q(\cdot)}$ is not equivalent to the unit vector basis in ℓ_2 for every $q \in \Omega(p)$ (otherwise we would arrive to contradiction by Theorem 3.3); therefore, we obtain

Corollary 3.1. *Suppose that a function p satisfies the conditions of Theorem 3.2(b). Then there exists a function $q \in \Omega(p)$ such that the Rademacher system is not equivalent in $L^{q(\cdot)}$ to the unit vector basis in ℓ_2.*

References

[Astashkin 2009] S. V. Astashkin, "Rademacher functions in symmetric spaces", *Sovrem. Mat. Fundam. Napravl.* **32** (2009), 3–161. In Russian; translated in *J. Math. Sci. (N.Y.)* **169**:6 (2010), 725–886. MR 2011f:46030 Zbl 1229.46020

[Astashkin and Maligranda 2010] S. V. Astashkin and L. Maligranda, "Rademacher functions in Cesàro type spaces", *Studia Math.* **198**:3 (2010), 235–247. MR 2011m:46040 Zbl 1202.46031

[Astashkin and Semënov 2013] S. V. Astashkin and E. M. Semënov, "Пространства, определяемые функцией Пэли", *Mat. Sb.* **204**:7 (2013), 3–24. Translated as "Spaces defined by the Paley function" in *Sbornik Math.* **204**:7 (2013), 937–957. MR 3114872 Zbl 1287.46021

[Astashkin et al. 2011] S. V. Astashkin, M. V. Leibov, and L. Maligranda, "Rademacher functions in BMO", *Studia Math.* **205**:1 (2011), 83–100. MR 2012h:46051 Zbl 1242.46034

[Bennett and Sharpley 1988] C. Bennett and R. Sharpley, *Interpolation of operators*, Pure and Applied Mathematics **129**, Academic Press, Boston, 1988. MR 89e:46001 Zbl 0647.46057

[Blei 2001] R. Blei, *Analysis in integer and fractional dimensions*, Cambridge Studies in Advanced Mathematics **71**, Cambridge University Press, 2001. MR 2003a:46008 Zbl 1006.46001

[Cruz-Uribe and Fiorenza 2013] D. V. Cruz-Uribe and A. Fiorenza, *Variable Lebesgue spaces: foundations and harmonic analysis*, Birkhäuser, Heidelberg, 2013. MR 3026953 Zbl 1268.46002

[Cruz-Uribe et al. 2014] D. V. Cruz-Uribe, A. Fiorenza, M. V. Ruzhansky, and J. Wirth, *Variable Lebesgue spaces and hyperbolic systems* (Barcelona, 2011), edited by S. Tikhonov, Birkhäuser, Basel, 2014. MR 3364250 Zbl 1297.46003

[Diening et al. 2011] L. Diening, P. Harjulehto, P. Hästö, and M. Růžička, *Lebesgue and Sobolev spaces with variable exponents*, Lecture Notes in Mathematics **2017**, Springer, Heidelberg, 2011. MR 2790542 Zbl 1222.46002

[Diestel et al. 1995] J. Diestel, H. Jarchow, and A. Tonge, *Absolutely summing operators*, Cambridge Studies in Advanced Mathematics **43**, Cambridge University Press, 1995. MR 96i:46001 Zbl 0855.47016

[Fiorenza and Rakotoson 2007] A. Fiorenza and J. M. Rakotoson, "Relative rearrangement and Lebesgue spaces $L^{p(\cdot)}$ with variable exponent", *J. Math. Pures Appl.* (9) **88**:6 (2007), 506–521. MR 2009b:46063 Zbl 1137.46016

[Johnson et al. 1979] W. B. Johnson, B. Maurey, G. Schechtman, and L. Tzafriri, *Symmetric structures in Banach spaces*, Mem. Amer. Math. Soc. **19**:217, American Mathematical Society, Providence, RI, 1979. MR 82j:46025 Zbl 0421.46023

[Kreĭn et al. 1982] S. G. Kreĭn, J. Ī. Petunīn, and E. M. Semënov, *Interpolation of linear operators*, Translations of Mathematical Monographs **54**, American Mathematical Society, Providence, RI, 1982. MR 84j:46103 Zbl 0493.46058

[Lindenstrauss and Tzafriri 1979] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces, II: Function spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete **97**, Springer, Berlin, 1979. MR 81c:46001 Zbl 0403.46022

[Lorentz 1951] G. G. Lorentz, "On the theory of spaces Λ", *Pacific J. Math.* **1** (1951), 411–429. MR 13,470c Zbl 0043.11302

[Maurey and Pisier 1976] B. Maurey and G. Pisier, "Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach", *Studia Math.* **58**:1 (1976), 45–90. MR 56 #1388 Zbl 0344.47014

[Musielak 1983] J. Musielak, *Orlicz spaces and modular spaces*, Lecture Notes in Mathematics **1034**, Springer, Berlin, 1983. MR 85m:46028 Zbl 0557.46020

[Pisier 1986] G. Pisier, *Factorization of linear operators and geometry of Banach spaces*, CBMS Regional Conference Series in Mathematics **60**, American Mathematical Society, Providence, RI, 1986. MR 88a:47020 Zbl 0588.46010

[Rodin and Semënov 1975] V. A. Rodin and E. M. Semënov, "Rademacher series in symmetric spaces", *Anal. Math.* **1**:3 (1975), 207–222. MR 52 #8905 Zbl 0315.46031

[Rodin and Semënov 1979] V. A. Rodin and E. M. Semënov, "О дополняемости подпространства, порожденного системой Радемахера, в симметричном пространстве", *Funktsional. Anal. i Prilozhen.* **13**:2 (1979), 91–92. Translated as "Complementability of the subspace generated by the Rademacher system in a symmetric space" in *Funct. Anal. Appl.* **13**:2 (1979), 150–151. MR 80j:46048 Zbl 0424.46025

[Sagher and Zhou 1990] Y. Sagher and K. C. Zhou, "A local version of a theorem of Khinchin", pp. 327–330 in *Analysis and partial differential equations*, edited by C. Sadosky, Lecture Notes in Pure and Applied Mathematics **122**, Dekker, New York, 1990. MR 91e:42039 Zbl 0694.42027

[Sharapudinov 1986] I. I. Sharapudinov, "О базисности системы Хаара в пространстве $\mathscr{L}^{p(t)}([0,1])$ и принципе локализации в среднем", *Mat. Sb.* (*N.S.*) **130(172)**:2 (1986), 275–283. Translated as "On the basis property of the Haar system in the space $\mathscr{L}^{p(t)}([0,1])$ and the principle of localization in the mean" in *Math. USSR Sbornik* **58**:1 (1987), 279–287. MR 88b:42034 Zbl 0639.42026

[Szarek 1976] S. J. Szarek, "On the best constants in the Khinchin inequality", *Studia Math.* **58**:2 (1976), 197–208. MR 55 #3672 Zbl 0424.42014

Received 27 Mar 2015. Revised 24 Jul 2015. Accepted 7 Sep 2015.

SERGEY ASTASHKIN: astash@samsu.ru
Department of Mathematics and Mechanics, Samara State University, Acad. Pavlov, 1, 443011 Samara, Russia

and

Samara State Aerospace University (SSAU), Moskovskoye shosse 34, 443086, Samara, Russia

MIECZYSŁAW MASTYŁO: mastylo@amu.edu.pl
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Collegium Mathematicum, Umultowska 87 Street 61-614 Poznań, Poland

and

Institute of Mathematics, Polish Academy of Sciences (Poznań branch), ul. Śniadeckich 8, 00-656 Warsaw, Poland

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.15

NONEXISTENCE OF SMALL DOUBLY PERIODIC SOLUTIONS FOR DISPERSIVE EQUATIONS

DAVID M. AMBROSE AND J. DOUGLAS WRIGHT

We study the question of existence of time-periodic, spatially periodic solutions for dispersive evolution equations, and in particular, we introduce a framework for demonstrating the nonexistence of such solutions. We formulate the problem so that doubly periodic solutions correspond to fixed points of a certain operator. We prove that this operator is locally contracting, for almost every temporal period, if the Duhamel integral associated to the evolution exhibits a weak smoothing property. This implies the nonexistence of nontrivial, small-amplitude time-periodic solutions for almost every period if the smoothing property holds. This can be viewed as a partial analogue of scattering for dispersive equations on periodic intervals, since scattering in free space implies the nonexistence of small coherent structures. We use a normal form to demonstrate the smoothing property on specific examples, so that it can be seen that there are indeed equations for which the hypotheses of the general theorem hold. The nonexistence result is thus established through the novel combination of small-divisor estimates and dispersive smoothing estimates. The examples treated include the Korteweg–de Vries equation and the Kawahara equation.

1. Introduction

In the absence of the ability to "explicitly" compute solutions of the Cauchy problem for a nonlinear dispersive system by some specialized technique particular to the equation at hand (such as complete integrability), coherent structures often form the backbone for both qualitative and quantitative descriptions of the dynamics of the system. Such structures, be they traveling waves, self-similar solutions, time-periodic solutions or some other sort of solution, give great insight into the short-time behavior of the system and often provide possible states towards which solutions trend as time goes to infinity.

For dispersive equations in free space, many authors have proved scattering results; we cannot hope to list all such results here, but a sampling is [Christ and Weinstein 1991; Ginibre and Velo 1986; Liu 1997; Ponce and Vega 1990; Strauss 1974]. For instance, Strauss [1974] showed that for a generalized Korteweg–de Vries (KdV) equation and for a nonlinear Schrödinger equation, all sufficiently small solutions decay to zero. There is no generally agreed upon meaning for scattering on periodic intervals, and one cannot expect

Ambrose gratefully acknowledges support from the National Science Foundation through grant DMS-1016267. Wright gratefully acknowledges support from the National Science Foundation through grant DMS-1105635.
MSC2010: 35B10, 35Q53.
Keywords: doubly periodic solutions, dispersive equations, small divisors, smoothing.

decay of solutions. However, decay of solutions implies the nonexistence of small-amplitude coherent structures, and the nonexistence of small coherent structures is a question which can be studied on periodic intervals. In the present work, we study the nonexistence of small time-periodic solutions for dispersive equations on periodic intervals. We prove a general theorem, showing that the existence of sufficiently strong dispersive smoothing effects implies the nonexistence of small doubly periodic solutions for almost every temporal period; we then demonstrate the required dispersive smoothing for particular examples. We have partially described the results of the present work briefly in the announcement [Ambrose and Wright 2014].

One method of constructing doubly periodic solutions of dispersive equations is to use Nash–Moser-type methods. These methods typically work in a function space with periodic boundary conditions in both space and time, so that the Fourier transform of the evolution equation can be taken in both variables. Then, a solution is sought nearby to an equilibrium solution. The implicit function theorem cannot be used due to the presence of small divisors. Instead, the small divisors are compensated for by the fast convergence afforded by Newton's method. A version of such arguments is now known as the Craig–Wayne–Bourgain method [Bourgain 1994; 1995a; Craig and Wayne 1993; Wayne 1990], which they used to demonstrate existence of doubly periodic solutions for a number of equations, such as nonlinear wave and nonlinear Schrödinger equations.

Such methods have since been extended to other systems, such as the irrotational gravity water wave, on finite or infinite depth, by Plotnikov, Toland, and Iooss [Plotnikov and Toland 2001; Iooss et al. 2005], or irrotational gravity-capillary waves by Alazard and Baldi [2015]. Also, Baldi [2013] used such methods to demonstrate existence of doubly periodic solutions for perturbations of the Benjamin–Ono equation. The typical result of these small-divisor methods is the existence of small-amplitude doubly periodic waves for the system under consideration, for certain values of the relevant parameters. One such parameter is the frequency (or equivalently, the temporal period) of the solution; other parameters may arise in specific applications, such as the surface tension parameter in [Alazard and Baldi 2015]. With these methods, the parameter values for which solutions are shown to exist are typically in a Cantor set.

The most classical version of small-divisor theory is due to Kolmogorov, Arnold, and Moser (KAM). KAM theory has been used by Kuksin [1988; 1998] to show that quasiperiodic solutions of the KdV equation persist under certain perturbations; see also the book of Kappeler and Pöschel [2003] and the references therein. Our work is complementary to these approaches, as we offer a nonexistence theory. That is, we give certain regions of frequency-amplitude space in which doubly periodic solutions cannot exist, whereas KAM theory gives instead regions where such solutions do in fact exist. Furthermore, we mention that the method of the present work requires no special structure on the equations; we only need certain estimates to be satisfied to conclude nonexistence of small doubly periodic solutions, and these are likely unrelated to any integrable or near-integrable structure.

For certain completely integrable equations, time-periodic waves can be shown to exist by producing exact, closed-form solutions. This is the case, for instance, for the KdV equation [Dubrovin 1975] and the Benjamin–Ono equation [Dobrokhotov and Krichever 1991; Matsuno 2004; Satsuma and Ishimori 1979]. The first author and Wilkening found that these time-periodic solutions of the Benjamin–Ono equation form continuous families [Ambrose and Wilkening 2009; 2010; Wilkening 2008]; this is in sharp contrast to results proved by small-divisor methods, then. The small-divisor methods do not address the question of whether such results are optimal; it is not possible at present to conclude whether or not time-periodic solutions exist as continuous families for such equations.

Thus, we find it of significant interest to develop further tools to answer the questions of both existence and nonexistence of doubly periodic waves for dispersive partial differential equations. In the present contribution, we develop a framework by which the nonexistence of small-amplitude time-periodic waves can be established. We do this by first using the Duhamel formula together with the time-periodic ansatz. Then we rewrite this formula, factoring out the linear operator. The resulting equation yields the notion of time-periodic solutions as fixed points of a new operator, which is the composition of the inverse of the linear operator with the Duhamel integral.

We then prove a general theorem, showing that under certain conditions, this operator is contracting in a neighborhood of the origin in a certain function space. Since the operator is the composition of two operators, we prove estimates for these individual operators separately. We are able to prove that the inverse of the linear operator acts like differentiation of order p for some $p > 1$. Of course, in order to have the contracting property, the composition must map from some function space, X, to itself. We use Sobolev spaces, so since the inverse of the linear operator acts like differentiation of order p, the Duhamel integral must satisfy an estimate with some form of a gain at least p derivatives. In Theorem 4, then, we have a general condition for the nonexistence of small-amplitude time-periodic waves for almost every possible temporal period: if the Duhamel integral possesses a weak form of smoothing (with an associated estimate), then the equation does not possess arbitrarily small doubly periodic waves for almost every period. The results described thus far are the content of Section 2. We mention that we are aware of one other result in the literature on nonexistence of small time-periodic solutions for almost every period; de la Llave [2000] uses a variational method to demonstrate nonexistence of small doubly periodic solutions for nonlinear wave equations. The methods of proof of the current work and of [de la Llave 2000] appear to be quite different, since in the present work we rely on smoothing estimates which derive from the linear part of the evolution; such smoothing estimates would not be expected to hold in the case of wave equations.

Clearly, we must address the question of whether there is an equation for which the truth of the hypotheses of Theorem 4 can be demonstrated. We demonstrate the required smoothing for the Duhamel integral associated to some dispersive evolution equations with fifth-order dispersion in Sections 3 and 4.1, and with seventh-order dispersion in Section 4.2. For these equations, the Duhamel integral satisfies a strong smoothing property: the Duhamel integral gains $p > 1$ derivatives, compensating for the loss of derivatives from the inverse of the linear operator. Theorem 4 does not, however, require so strong a smoothing property. In Section 5, we demonstrate that the Duhamel integral for the KdV equation satisfies a weaker smoothing property, allowing Theorem 4 to be applied and demonstrating the nonexistence of small doubly periodic solutions for almost every temporal period.

The estimates for the inverse of the linear operator (Lemma 1, Lemma 2, and Corollary 3 below), in which we demonstrate that the inverse of our linear operator acts like differentiation of order p, are proved by small-divisor techniques. In fact, these are versions of classical results, such as can be found, for instance, in [Ghys 2007]. As with all small-divisor results, some parameter values are discarded;

in the present case, the parameter is the temporal period of the solution. Thus, we arrive at a result about nonexistence of small solutions for almost every possible temporal period. Even though such a small-divisor argument is classical, we provide our own proof because the detailed information about the set of temporal periods in the proof is helpful.

For our particular examples of dispersive equations, we prove smoothing estimates for the Duhamel integral by following the lines of an argument by Erdoğan and Tzirakis [2013]. By using a normal form representation, they showed that the Duhamel integral for the KdV equation gains $1-\varepsilon$ derivatives as compared to the initial data, for any $\varepsilon > 0$. (We mention that a similar result has been demonstrated on the real line by Linares and Scialom [1993].) For the equations with fifth-order dispersion which we consider in Sections 3 and 4.1, we find a gain of two derivatives. This is in line with the usual, expected gain of regularity from dispersion. In general, if the dispersion relation is of order r (say, for instance, the linearized evolution equation is, in the Fourier transform, $\hat{w}_t = ik^r\hat{w}$), then one expects to gain $(r-1)/2$ derivatives in some sense [Kenig et al. 1991]; this is known as the Kato smoothing effect [1983]. With fifth-order dispersion, this means the expected gain is two derivatives. Given the results of [Erdoğan and Tzirakis 2013] as well as the present work, it does appear that it is reasonable to expect the same order of smoothing in the spatially periodic setting, for the Duhamel integral. In fact, in Section 4.2, we show that for an evolution equation with seventh-order dispersion, the gain of regularity on the Duhamel integral is four derivatives; thus, the smoothing effect in [Erdoğan and Tzirakis 2013] and the present work is not the same as Kato smoothing, but is still due to the presence of dispersion. To demonstrate our weaker smoothing estimate for the KdV equation in Section 5, we begin with the same normal form as before, but we estimate the terms differently.

We close with some discussion in Section 6.

2. Nonexistence of doubly periodic solutions

We begin with the evolution equation

$$\partial_t u = Au + Nu, \tag{1}$$

where A is a linear operator and N is a nonlinear operator. Then, the solution of (1) with initial data $u(\cdot,0) = u_0$, if there is a solution, can be represented with the usual Duhamel formula,

$$u(\cdot,t) = e^{At}u_0 + \int_0^t e^{A(t-\tau)}N(u(\cdot,\tau))\,d\tau. \tag{2}$$

Given a time t, we define the linear solution operator $S_L(t) = e^{At}$ and the difference of the solution operator and the linear solution operator to be $S_D(t)$; thus, S_D is exactly the Duhamel integral:

$$S_D(t)u_0 = \int_0^t e^{A(t-\tau)}N(u(\cdot,\tau))\,d\tau.$$

We work in the spatially periodic case, so we assume that solutions u of (1) satisfy

$$u(x+2\pi,t) = u(x,t) \quad \forall x \in \mathbb{R}.$$

We assume that (1) maintains the mean of solutions; that is, given any u in a reasonable function space, we have

$$\int_0^{2\pi} u(x,t)\,dx = \int_0^{2\pi} u_0(x)\,dx.$$

For the remainder of the present section, we will assume that u_0 (and thus $u(\cdot\,,t)$) has mean zero.

If u_0 is the initial data for a time-periodic solution of (1) with temporal period T, then we have

$$u_0 = S_L(T)u_0 + S_D(T)u_0.$$

We rewrite this as

$$\big(I - S_L(T) - S_D(T)\big)u_0 = 0. \tag{3}$$

Our goal is to demonstrate nonexistence of small-amplitude doubly periodic solutions of (3) for certain temporal periods. We begin now to focus only on certain values of T, and our first restriction on values of T is to ensure that $I - S_L(T)$ is invertible. For $s > 0$, define the space H_0^s to be the subset of the usual spatially periodic Sobolev space H^s such that for all $f \in H_0^s$, the mean of f is equal to zero. We assume that the operator $S_L(t)$ is bounded,

$$S_L(t) : H_0^s \to H_0^s \quad \forall t \in \mathbb{R}.$$

Then, we define the set W to be

$$W = \big\{t \in (0,\infty) : \ker(I - S_L(t)) = \{0\}\big\}.$$

For any $T \in W$, we rewrite (3) by factoring out $I - S_L(T)$:

$$(I - S_L(T))\big(I - (I - S_L(T))^{-1} S_D(T)\big)u_0 = 0.$$

We see, then, that if u_0 is the initial data for a time-periodic solution of (1) with temporal period $T \in W$, then u_0 is a fixed point of the operator

$$K(T) := (I - S_L(T))^{-1} S_D(T).$$

If we can show that this is (locally) a contraction on H_0^s, then there are no (small) nontrivial time-periodic solutions in the space H_0^s with temporal period T. To establish this, we will need estimates both for $(I - S_L(T))^{-1}$ and for $S_D(T)$. In Sections 2.1 and 2.2, we establish estimates for $(I - S_L(T))^{-1}$; the results are that the symbol can be bounded as $|k|^p$, where k is the variable in Fourier space for some $p > 1$, for certain values of T. Thus, the inverse of the linear operator acts like differentiation of order $p > 1$. In Section 2.3, then, we will state a corollary of these estimates: if the operator S_D satisfies a certain estimate related to a gain of p derivatives, then the operator $K(T)$ is locally contracting, and thus there are no small time-periodic solutions with temporal period T.

2.1. ***The linear estimate: the homogeneous case.*** We now prove our estimate for $(I - S_L(T))^{-1}$ in the case that the linear operator A has symbol

$$\mathcal{F}(A)(k) = ik^r \quad \forall k \in \mathbb{Z}, \tag{4}$$

with r being an odd integer. This estimate is the content of Lemma 1. We note that this is not, strictly speaking, useful to us, as we will prove a version of the lemma for a more general class of operators A in the following section. However, we include Lemma 1 because the simplicity of the form of (4) allows for clarity of exposition. The proof of Lemma 2 below will build upon the proof of Lemma 1, which we now present.

We mention that the restriction that r is an odd integer is for two reasons, neither of which is essential. First, since we are interested in equations of KdV-type, these are the natural values of r to consider. Second, the restriction is for ease of exposition, since Lemma 1 would be equally true if we allowed the symbol of A to be defined by

$$\mathcal{F}(A)(k) = i|k|^{r_1} k^{r_2} \tag{5}$$

for all $k \in \mathbb{Z}$, with $r_1 \in (0, \infty)$ and $r_2 \in \{0, 1\}$. The presentation, however, is slightly streamlined if we use (4) instead of (5).

Lemma 1. *Let the linear operator A be given by* (4). *Let $0 < T_1 < T_2$ be given. Let $0 < \delta < T_2 - T_1$ be given. Let $p > 1$ be given. There exists a set $W_{p,\delta} \subseteq [T_1, T_2] \cap W$ and there exists $c_1 > 0$ such that the Lebesgue measure of $W_{p,\delta}$ satisfies $\mu(W_{p,\delta}) > T_2 - T_1 - \delta$, and for all $k \in \mathbb{Z} \setminus \{0\}$, for all $T \in W_{p,\delta}$, we have*

$$\left|\mathcal{F}(I - S_L(T))^{-1}(k)\right| < c_1|k|^p.$$

Proof. For the moment, we fix $k \in \mathbb{Z} \setminus \{0\}$. We need to estimate

$$\left|\mathcal{F}(I - S_L(T))^{-1}(k)\right| = \left|\frac{1}{1 - \exp\{ik^r T\}}\right| = \frac{1}{\sqrt{2}}(1 - \cos(k^r T))^{-1/2}. \tag{6}$$

Clearly, the symbol of the inverse operator is undefined if there exists $n \in \mathbb{N}$ such that $T = 2\pi n/|k|^r$. With the assumption that $T \in [T_1, T_2]$, the associated values of n comprise the set

$$\mathcal{N} := \left[\frac{|k|^r T_1}{2\pi}, \frac{|k|^r T_2}{2\pi}\right] \cap \mathbb{N}.$$

We remove a small set of possible periods around each of these values; that is to say, we consider

$$T \in [T_1, T_2] \setminus \bigcup_{n \in \mathcal{N}} \left[\frac{2\pi n}{|k|^r} - \varepsilon, \frac{2\pi n}{|k|^r} + \varepsilon\right] \tag{7}$$

for some $0 < \varepsilon \ll 1$ to be specified.

To start, we may notice that the collection of intervals $\left[2\pi n/|k|^r - \varepsilon, 2\pi n/|k|^r + \varepsilon\right]$ do not overlap for different values of n as long as, for all $n \in \mathcal{N}$, we have

$$\frac{2\pi n}{|k|^r} + \varepsilon < \frac{2\pi(n+1)}{|k|^r} - \varepsilon.$$

This is satisfied as long as $\varepsilon < \pi/|k|^r$. When we choose ε, this condition will be satisfied.

A simple calculation shows that, on an interval of values of θ which does not include an integer multiple of 2π, the value of $(1 - \cos\theta)$ is minimized at the endpoints of the interval. Thus, the value of (6) is largest on our set of possible periods at the values $T = 2\pi n/|k|^r \pm \varepsilon$ for $n \in \mathcal{N}$. At such values, we find

$$\left|\mathcal{F}\left(\left(I - S_L\left(\frac{2\pi n}{|k|^r} \pm \varepsilon\right)\right)^{-1}\right)(k)\right| = \frac{1}{\sqrt{2}\left(1-\cos\left(k^r\left(\frac{2\pi n}{|k|^r} \pm \varepsilon\right)\right)\right)^{1/2}}$$

$$= \frac{1}{\sqrt{2}(1-\cos(\pm 2\pi n \pm \varepsilon k^r))^{1/2}} = \frac{1}{\sqrt{2}(1-\cos(\varepsilon k^r))^{1/2}}. \quad (8)$$

We will now perform a Taylor expansion for cosine, paying attention to the error estimates. For any $\theta \in \mathbb{R}$, we have the formula

$$\cos\theta = 1 - \tfrac{1}{2}\theta^2 + \tfrac{1}{6}\sin(\xi)\theta^3$$

for some ξ between 0 and θ. We notice that $\left|\frac{1}{6}\sin(\xi)\theta^3\right| \le \frac{1}{4}\theta^2$, as long as $|\theta| \le \frac{3}{2}$. Thus, for $|\theta| \le \frac{3}{2}$, we have

$$|1-\cos\theta| = \left|\tfrac{1}{2}\theta^2 - \tfrac{1}{6}\sin(\xi)\theta^3\right| \ge \tfrac{1}{4}\theta^2. \quad (9)$$

Combining (9) with (8), we find that for any T satisfying (7), if $\varepsilon|k|^r < \frac{3}{2}$, then we have

$$\left|\mathcal{F}\left((I - S_L(T))^{-1}(k)\right)\right| \le \frac{1}{(\sqrt{2})\left(\frac{1}{2}\varepsilon|k|^r\right)} = \frac{\sqrt{2}}{\varepsilon|k|^r}. \quad (10)$$

We choose $\varepsilon = c_0|k|^{-p-r}$. This choice of ε immediately yields the claimed estimate for the symbol. Recall that we have specified $p > 1$; the positive constant c_0 is to be specified. The conditions we have placed on ε above are (1) $\varepsilon < \pi/|k|^r$ and (2) $\varepsilon|k|^r < \frac{3}{2}$. These conditions are both satisfied as long as $c_0 < \frac{3}{2}$.

For fixed $k \in \mathbb{Z}\setminus\{0\}$ and for fixed $n \in \mathcal{N}$, we have removed a set of measure $2\varepsilon = 2c_0|k|^{-r-p}$ from the interval $[T_1, T_2]$. Since $\mathcal{N}$ is the intersection of an interval with the natural numbers, we see that the cardinality of $\mathcal{N}$ is less than $|k|^r(T_2 - T_1) + 1$, so for fixed k, we have removed a set of measure no more than $2c_0(1+T_2-T_1)|k|^{-p}$. Summing over k, since we have chosen $p > 1$, the measure of the set we have removed is finite. Taking c_0 sufficiently small, we can conclude that the measure of the set which is removed has Lebesgue measure smaller than δ. To be definite, we write the definition of the set $W_{p,\delta}$, which is

$$W_{p,\delta} = [T_1, T_2] \setminus \bigcup_{k\in\mathbb{Z}\setminus\{0\}} \bigcup_{n\in\mathcal{N}} \left[\frac{2\pi n}{|k|^r} - \frac{c_0}{|k|^{r+p}}, \frac{2\pi n}{|k|^r} + \frac{c_0}{|k|^{r+p}}\right],$$

where c_0 is chosen so that $0 < c_0 < \frac{3}{2}$, and also so that

$$c_0 < \frac{\delta}{2(1+T_2-T_1)\sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-p}}.$$

Finally, the constant c_1 is given by $c_1 = \sqrt{2}/c_0$. This completes the proof. □

We note that we can see clearly the dependence of the constant c_1 on the parameters T_1, T_2, p, and δ. If we want a larger set of potential periods, we could take $T_2 - T_1$ larger or δ smaller; this would result in a larger value of c_1. Choosing smaller values of $p > 1$ also leads to larger values of c_1.

2.2. ***The linear estimate: the nonresonant case.*** The estimate of Section 2.1 can be generalized to allow operators which include lower-order terms, in some cases: there must not be a resonance between the different terms, in the sense that we require $\mathcal{F}A(k) \neq 0$ for any nonzero $k \in \mathbb{Z}$. To be very precise, we consider linear operators A which satisfy the conditions **(H)**, which we now describe.

(**H**) Let $M \in \mathbb{N}$ be given, with $M \geq 2$. For each $m \in \{1, 2, \ldots, M\}$, let r_m be an odd integer, with these r_m satisfying $r_1 > r_2 > \cdots > r_M > 0$. For all $m \in \{1, 2, \ldots, M\}$, let $Z_m \subseteq \mathbb{R}$ be bounded. Let $Z = Z_1 \times Z_2 \times \cdots \times Z_M$. Let $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_M)$. Assume there exists $\beta_1 > 0$ and $\beta_2 > 0$ such that for all $\alpha_1 \in Z_1$,

$$|\alpha_1| > \beta_1, \tag{11}$$

and for all $\vec{\alpha} \in Z$, for all $k \in \mathbb{Z} \setminus \{0\}$,

$$\left| \sum_{m=1}^{M} \alpha_m k^{r_m} \right| \geq \beta_2. \tag{12}$$

Given $\vec{\alpha} \in Z$, let the linear operator A be defined through its symbol as

$$\mathcal{F}(A)(k) = i \sum_{m=1}^{M} \alpha_m k^{r_m} \quad \forall k \in \mathbb{Z}. \tag{13}$$

We again remark that the condition that the r_m be odd integers is for ease of exposition and because these are the relevant values for the KdV equation and for the other equations (such as Kawahara equations) which we consider in the present work; as mentioned previously, this restriction can be relaxed with no difficulty. We further remark that the condition (11) ensures that the equation is dispersive of order r_1 (i.e., the leading-order term is of the same order for all $\vec{\alpha} \in Z$). The condition (12) ensures that the symbol never vanishes.

Lemma 2. *Let the set Z and the linear operator A satisfy the hypotheses* (**H**), *so that in particular*, A *is defined by* (13). *Let $0 < T_1 < T_2$ be given. Let $0 < \delta < T_2 - T_1$ be given. Let $p > 1$ be given. There exists a set $W_{p,\delta} \subseteq [T_1, T_2] \cap W$ and there exists $c_1 > 0$ such that the Lebesgue measure of $W_{p,\delta}$ satisfies $\mu(W_{p,\delta}) > T_2 - T_1 - \delta$, and such that for all $\vec{\alpha} \in Z$, for all $k \in \mathbb{Z} \setminus \{0\}$, for all $T \in W_{p,\delta}$, we have*

$$\left| \mathcal{F}(I - S_L(T))^{-1}(k) \right| < c_1 |k|^p.$$

Proof. The proof of Lemma 1 can be repeated, with k^r replaced in every instance by $\sum_{m=1}^{M} \alpha_m k^{r_m}$, until (10); to be concise, we introduce the notation

$$\lambda_k = \sum_{m=1}^{M} \alpha_m k^{r_m}.$$

We choose $\varepsilon = c_0 |k|^{-p-r_1}$, with c_0 to be specified.

Similarly to the previous case, we have the conditions (i) $\varepsilon < \pi |\lambda_k|^{-1}$, and (ii) $\varepsilon |\lambda_k| < \frac{3}{2}$. The relevant product satisfies

$$|\varepsilon \lambda_k| = c_0 \left| \sum_{m=1}^{M} \alpha_m k^{r_m} |k|^{-r_1-p} \right| \leq c_0 \sum_{m=1}^{M} |\alpha_m| \quad \forall k \in \mathbb{Z} \setminus \{0\}.$$

Thus, recalling that the sets Z_m are bounded, we can clearly take c_0 sufficiently small to satisfy conditions (i) and (ii).

We must revisit the estimate (10) in the present setting. We have

$$\left|\mathcal{F}(I-S_L(T))^{-1}(k)\right| \leq \frac{\sqrt{2}}{\varepsilon|\lambda_k|} = \frac{\sqrt{2}}{c_0}\frac{|k|^p}{|\lambda_k k^{-r_1}|}.$$

Since λ_k is never equal to zero, we conclude that $\lambda_k k^{-r_1}$ is also never equal to zero for $k \in \mathbb{Z}\setminus\{0\}$. Furthermore, there exists $K \in \mathbb{N}$ such that for all $k \in \mathbb{Z}$ satisfying $|k| > K$, for all $\vec{\alpha} \in Z$, we have

$$\left|\sum_{m=2}^{M} \alpha_m k^{r_m - r_1}\right| \leq |K|^{r_2-r_1} \sum_{m=2}^{M} |\alpha_m| < \tfrac{1}{2}\beta_1.$$

This implies that for all $k \in \mathbb{Z}$ with $|k| > K$, for all $\vec{\alpha} \in Z$,

$$|\lambda_k k^{-r_1}| = \left|\alpha_1 + \sum_{m=2}^{\infty} \alpha_m k^{r_m-r_1}\right| \geq \tfrac{1}{2}\beta_1 > 0,$$

$$|\lambda_k k^{-r_1}| = \left|\alpha_1 + \sum_{m=2}^{\infty} \alpha_m k^{r_m-r_1}\right| \leq \big(\sup_{\alpha_1\in Z_1} |\alpha_1|\big) + \tfrac{1}{2}\beta_1.$$

Furthermore, for any $k \in \mathbb{Z}\setminus\{0\}$ satisfying $|k| \leq K$, we know

$$\inf_{\vec{\alpha}\in Z} \frac{|\lambda_k|}{|k^{r_1}|} = \inf_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}| \in (0,\infty),$$

$$\sup_{\vec{\alpha}\in Z} \frac{|\lambda_k|}{|k^{r_1}|} = \sup_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}| \in (0,\infty).$$

Combining this information, we see that

$$\inf_{k\in\mathbb{Z}\setminus\{0\}} \inf_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}| \in (0,\infty),$$

$$\sup_{k\in\mathbb{Z}\setminus\{0\}} \sup_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}| \in (0,\infty).$$

Our value of c_1 is therefore

$$c_1 = \frac{\sqrt{2}}{c_0}\big(\inf_{k\in\mathbb{Z}\setminus\{0\}} \inf_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}|\big)^{-1}.$$

In the current setting, for each $\vec{\alpha} \in Z$, for each $k \in \mathbb{Z}\setminus\{0\}$, the set $\mathcal{N}$ is defined by

$$\mathcal{N} = \left[\frac{T_1|\lambda_k|}{2\pi}, \frac{T_2|\lambda_k|}{2\pi}\right] \cap \mathbb{N}.$$

Thus, for all $\vec{\alpha} \in Z$, for all $k \in \mathbb{Z}\setminus\{0\}$, the cardinality of $\mathcal{N}$ satisfies

$$\begin{aligned}\operatorname{card}(\mathcal{N}) &\leq (T_2-T_1)|\lambda_k| + 1 \leq (T_2-T_1)|k|^{r_1}|\lambda_k k^{-r_1}| + 1\\ &\leq (T_2-T_1)|k|^{r_1}\big(\sup_{k\in\mathbb{Z}\setminus\{0\}} \sup_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}|\big) + 1.\end{aligned}$$

We then take the product $2\varepsilon\,\mathrm{card}(\mathcal{N})$, finding the estimate

$$2\varepsilon\,\mathrm{card}(\mathcal{N}) \leq 2c_0\big(1 + (T_2 - T_1) \sup_{k\in\mathbb{Z}\setminus\{0\}} \sup_{\vec{\alpha}\in Z} |\lambda_k k^{-r_1}|\big)|k|^{-p}.$$

As in the proof of Lemma 1, we sum over k, and we find that we can take c_0 sufficiently small to satisfy the remaining claims of the lemma. □

Lemma 2 allows for two kinds of uniformity: the same set $W_{p,\delta}$ works for all $\vec{\alpha} \in Z$, and the constant c_1 is able to be used for all $T \in W_{p,\delta}$. The cost of this uniformity with respect to the constant c_1 is that the set $W_{p,\delta}$ does not have full measure. By sending δ to zero, we can achieve an estimate for almost every $T \in [T_1, T_2]$, but then the constant will depend on the choice of T. In doing this, we are able to maintain the uniformity with respect to the set Z. This is the content of the following corollary.

Corollary 3. *Let the set Z and the linear operator A satisfy the hypotheses* (**H**), *with A given by* (13). *Let $p > 1$ be given. Let $0 < T_1 < T_2$ be given. For almost every $T \in [T_1, T_2]$, there exists $c > 0$ such that for all $\vec{\alpha} \in Z$ and for all $k \in \mathbb{Z} \setminus \{0\}$, we have the estimate*

$$\big|\mathcal{F}\big((I - S_L(T))^{-1}\big)(k)\big| \leq c|k|^p.$$

Proof. For any δ satisfying $0 < \delta < T_2 - T_1$, let the set $W_{p,\delta}$ be as in Lemma 2. For any $T \in [T_1, T_2]$, if there exists a value of δ such that $T \in W_{p,\delta}$, then the desired estimate is satisfied. Let

$$W_p = \bigcup_{n=2}^{\infty} W_{p,(T_2-T_1)/n}.$$

Then, for all $T \in W_p$, the estimate holds. Since for all $n \geq 2$ we have $W_{p,(T_2-T_1)/n} \subseteq W_p \subseteq [T_1, T_2]$, and since $\mu(W_{p,(T_2-T_1)/n}) \geq (T_2-T_1)(1-1/n)$, we see that $\mu(W_p) = T_2-T_1$. This completes the proof. □

2.3. ***The general theorem.*** We are now able to state a general theorem which follows from the above discussion. In Theorem 4, when we say that N is "as above", this includes the property that the evolution equation (1) preserves the mean value of the initial data. The theorem contains two statements about nonexistence of small-amplitude time-periodic solutions. The first statement is for a given $\delta > 0$; for T in the set $W_{p,\delta}$, we conclude that there is a uniform threshold for the amplitude of time-periodic solutions. For the second statement, we conclude that for almost any $T \in [T_1, T_2]$, there is a threshold for the amplitude of time-periodic solutions; this second statement is not uniform. These results are conditional on the existence of smoothing estimates. In Sections 3, 4.1, 4.2, and 5, we will demonstrate the required smoothing estimate for particular equations.

Theorem 4. *Let $0 < T_1 < T_2$ be given, and let $0 < \delta < T_2 - T_1$ be given. Let the set Z and the operator A satisfy the hypotheses* (**H**), *with A given by* (13). *Let the nonlinear operator N be as above. Assume there exists $p > 1$, $\tilde{p} \geq 0$, $q > 0$, $s \geq 0$, $c > 0$, and $\eta > 0$ such that for all $u_0 \in H_0^{s+\tilde{p}}$ with $\|u_0\|_{H^{s+\tilde{p}}} \leq \eta$, for all $\vec{\alpha} \in Z$, the following estimate is satisfied*:

$$\|S_D(T)u_0\|_{H^{s+p}} \leq c\|u_0\|_{H^s}\|u_0\|^q_{H^{s+\tilde{p}}} \tag{14}$$

for all $T \in W_{p,\delta}$. Then, there exists $r_0 > 0$ such that if u is a smooth, nontrivial, mean-zero time-periodic solution of (1) *with temporal period $T \in W_{p,\delta}$, then*

$$\inf_{t\in[0,T]} \|u\|_{H^{s+\tilde{p}}} \geq r_0. \tag{15}$$

Furthermore, if (14) *holds for every $T \in W_p$, then for every $T \in W_p$, there exists $r_0 > 0$ such that if u is a nontrivial mean-zero time-periodic solution of* (1) *with temporal period T, then* (15) *holds.*

Proof. The assumptions of the theorem, together with either Lemma 2 or Corollary 3, imply that there exists $C > 0$ such that for all $u_0 \in H_0^{s+\tilde{p}}$ satisfying $\|u_0\|_{H^{s+\tilde{p}}} < \eta$,

$$\|K(T)u_0\|_{H^s} \leq C\|u_0\|_{H^s}\|u_0\|^q_{H^{s+\tilde{p}}}.$$

If u_0 satisfies

$$0 < C\|u_0\|^q_{H^{s+\tilde{p}}} < 1,$$

then $\|K(T)u_0\|_{H^s} < \|u_0\|_{H^s}$, and thus u_0 cannot be a fixed point of $K(T)$. Thus, the only fixed point in a ball around zero is zero. ☐

Remark 5. In the announcement [Ambrose and Wright 2014], the version of this theorem which appeared was restricted to the case $\tilde{p} = 0$. In this case, the inverse of the linear operator acts like differentiation of order $p > 1$, and the Duhamel integral has a compensating gain of p derivatives. In Sections 3 and 4, we will give examples for which this smoothing property holds; these examples include the Kawahara equation. However, as the estimate (14) shows, what is needed is actually much weaker than the Duhamel integral gaining p derivatives; instead, it is only necessary that $s + p$ derivatives of the Duhamel integral satisfy a nonlinear estimate in which one factor involves only s derivatives. In Section 5, we will demonstrate that the estimate (14) holds with $\tilde{p} > 0$ for the KdV equation.

3. Application to a fifth-order dispersive equation

In this section, we will apply the above results to a specific dispersive equation, with sufficiently strong dispersion, with $\tilde{p} = 0$. We are using a version of the Erdoğan–Tzirakis argument [2013] for the equation

$$\partial_t \tilde{u} = \partial_x^5 \tilde{u} - 2\tilde{u}\,\partial_x \tilde{u} + \tilde{\omega}\,\partial_x \tilde{u} \tag{16}$$

(for any $\tilde{\omega} \in \mathbb{R}$) to get the desired smoothing effect. As we have discussed, we consider the spatially periodic case, with spatial period equal to 2π. We first rewrite (16) to remove the mean, and also to remove the tildes.

We consider the initial condition $\tilde{u}(x, 0) = \tilde{g}(x)$. Assume the mean of $\tilde{g}$ is equal to $\bar{g}$, which can be any real number. Let $u = \tilde{u} - \bar{g}$. Since the evolution for $\tilde{u}$ conserves the mean, the mean of u will equal zero at all times. The evolution equation satisfied by u is

$$\partial_t u = \partial_x^5 u - 2u\,\partial_x u + \omega\,\partial_x u, \tag{17}$$

where $\omega = \tilde{\omega} - 2\bar{g}$. The initial data for (17) is $g = \tilde{g} - \bar{g}$, which of course has mean zero.

We now discuss the appropriate existence theory for (17).

3.1. ***Existence theory.*** The well-posedness of the initial value problem for (17) (or for (16)) has been established in [Bourgain 1995b], in the space H^s for $s > 0$; a more general family of equations including these was also shown to be well-posed in H^s for $s > \frac{1}{2}$ in [Hu and Li 2015]. In the present work, we are not concerned with demonstrating results at the lowest possible regularity, but instead in finding estimates which will work with the nonexistence argument. Towards this end, we will give a simple existence theorem for the initial value problem for (17) in the space H^6, as the resulting estimates will be useful. We mention that the choice of H^6 as the function space is made so that the solutions are classical solutions.

Proposition 6. *Let $u_0 \in H^6$ be given. There exists $T > 0$ and a unique $u \in C([0,T]; H^6)$ such that u solves the initial value problem* (17) *with data $u(\cdot, 0) = u_0$.*

Proof. We begin by introducing a mollifier, $\mathcal{J}_\varepsilon$, for $\varepsilon > 0$. We use the mollifier to make an approximate evolution equation,

$$\partial_t u^\varepsilon = \mathcal{J}_\varepsilon^2\, \partial_x^5 u^\varepsilon - \mathcal{J}_\varepsilon(2(\mathcal{J}_\varepsilon u^\varepsilon)(\mathcal{J}_\varepsilon u_x^\varepsilon)) + \omega \mathcal{J}_\varepsilon^2 u_x^\varepsilon. \tag{18}$$

When combined with the mollifier, all of the derivatives on the right-hand side have become bounded operators, and thus solutions for the initial value problem for u^ε, with initial data $u^\varepsilon(\cdot, 0) = g \in H^6$, exist in $C([0, T_\varepsilon]; H^6)$ by Picard's theorem (see Chapter 3 of [Majda and Bertozzi 2002]).

In order to show that the interval of existence can be taken to be independent of ε, we must make an energy estimate. We let the energy functional be an equivalent version of the square of the H^6 norm:

$$\mathcal{E}(t) = \frac{1}{2}\int_0^{2\pi} (u^\varepsilon)^2 + (\partial_x^6 u^\varepsilon)^2\, dx.$$

Taking the time derivative, we find

$$\frac{d\mathcal{E}}{dt} = \int_0^{2\pi} (u^\varepsilon)(\partial_t u^\varepsilon)\, dx + \int_0^{2\pi} (\partial_x^6 u^\varepsilon)(\partial_t \partial_x^6 u^\varepsilon)\, dx = \mathrm{I} + \mathrm{II}.$$

For I, we plug in from the evolution equation, using the fact that $\mathcal{J}_\varepsilon$ is self-adjoint and commutes with ∂_x:

$$\mathrm{I} = \int_0^{2\pi} (\mathcal{J}_\varepsilon u^\varepsilon)\, \partial_x^5(\mathcal{J}_\varepsilon u^\varepsilon) - 2(\mathcal{J}_\varepsilon u^\varepsilon)^2(\mathcal{J}_\varepsilon u_x^\varepsilon) + \omega(\mathcal{J} u^\varepsilon)(\mathcal{J}_\varepsilon u_x^\varepsilon)\, dx.$$

All of these terms vanish upon integrating by parts and/or recognizing perfect derivatives; therefore, $\mathrm{I} = 0$.

To study the term II, it is helpful to first apply six spatial derivatives to (18). We use the product rule, finding

$$\partial_t \partial_x^6 u^\varepsilon = \mathcal{J}_\varepsilon^2 \partial_x^{11} u^\varepsilon - \mathcal{J}_\varepsilon(2(\mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^7 \mathcal{J}_\varepsilon u^\varepsilon)) + \omega \mathcal{J}_\varepsilon^2 \partial_x^7 u^\varepsilon - 2\mathcal{J}_\varepsilon\Bigg(\sum_{m=1}^{6} \binom{6}{m} (\partial_x^m \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^{7-m} \mathcal{J}_\varepsilon u^\varepsilon)\Bigg). \tag{19}$$

We can then write

$$\mathrm{II} = \mathrm{II}_1 + \mathrm{II}_2 + \mathrm{II}_3 + \mathrm{II}_4,$$

where each of these terms corresponds to one of the four terms on the right-hand side of (19). We will now deal with these one at a time.

We again will frequently use the fact that $\mathcal{J}_\varepsilon$ is self-adjoint and commutes with ∂_x. To begin, we have

$$\mathrm{II}_1 = \int_0^{2\pi} (\partial_x^6 \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^{11} \mathcal{J}_\varepsilon u^\varepsilon)\, dx.$$

After integrating by parts and recognizing a perfect derivative, we see that $\mathrm{II}_1 = 0$.

For II_2, we have

$$\mathrm{II}_2 = -2 \int_0^{2\pi} (\mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^6 \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^7 \mathcal{J}_\varepsilon u^\varepsilon)\, dx.$$

We recognize a perfect derivative and integrate by parts, finding

$$\mathrm{II}_2 = \int_0^{2\pi} (\mathcal{J}_\varepsilon u_x^\varepsilon)(\partial_x^6 \mathcal{J}_\varepsilon u^\varepsilon)^2\, dx \le c\mathcal{E}^{3/2}.$$

For II_3, we have

$$\mathrm{II}_3 = \omega \int_0^{2\pi} (\partial_x^6 \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^7 \mathcal{J}_\varepsilon u^\varepsilon)\, dx = 0.$$

As before, this integrates to zero because it is the integral of a perfect derivative over a periodic interval.

Finally, we treat II_4. We have

$$\mathrm{II}_4 = -2 \sum_{m=1}^{6} \binom{6}{m} \int_0^{2\pi} (\partial_x^6 \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^m \mathcal{J}_\varepsilon u^\varepsilon)(\partial_x^{7-m} \mathcal{J}_\varepsilon u^\varepsilon)\, dx. \tag{20}$$

Each of the integrals on the right-hand side of (20) can be bounded by $\mathcal{E}^{3/2}$.

We therefore conclude that there exists $c > 0$ such that

$$\frac{d\mathcal{E}}{dt} \le c\mathcal{E}^{3/2}.$$

This can be rephrased as

$$\frac{d\|u^\varepsilon\|_{H^6}}{dt} \le c\|u^\varepsilon\|_{H^6}^2.$$

This estimate clearly indicates that the solutions u^ε cannot blow up arbitrarily quickly, and thus exist on a common time interval. So, we have shown that there exists $T > 0$, independent of ε, such that for all $\varepsilon > 0$, we have $u^\varepsilon \in C([0, T]; H^6)$, with the norm bounded independently of ε.

Since u^ε is bounded uniformly (in both t and ε) and since u^ε solves (18), we see that $\partial_t u^\varepsilon$ is uniformly bounded in H^1, and thus in L^∞. By the Arzelà–Ascoli theorem, there exists a subsequence (which we do not relabel) and a limit, $u \in C([0, 2\pi] \times [0, T])$, such that $u_\varepsilon \to u$ in this space. Standard arguments (again, see Chapter 3 of [Majda and Bertozzi 2002], for instance) then imply that u belongs to the space $C([0, T]; H^6)$, that u obeys the same uniform bound as the u^ε, and that u is a solution of (17).

Uniqueness of solutions (and, in fact, continuous dependence on the initial data) follows from a more elementary version of the energy estimate. If we let $u \in C([0, T]; H^6)$ be a solution corresponding to initial data $u_0 \in H^6$, and if we let $v \in C([0, T]; H^6)$ be a solution corresponding to initial data $v_0 \in H^6$, then we can estimate the norm of $u-v$. If $E_d = \|u-v\|_{L^2}^2$, then a straightforward calculation, together with the uniform bounds established previously, indicates $\frac{d}{dt} E_d \le cE_d$. This implies that $E_d \le E_d(0)e^{ct}$ for all t for which the solutions are defined. If $u_0 = v_0$, then we see that $u = v$. This is the desired uniqueness result. □

Remark 7. These estimates are uniform in ω, since all of the terms in the energy estimate which involve ω are equal to zero.

3.2. ***Reformulation.*** In this section, we use a normal form [Shatah 1985] to rewrite the evolution equation in a beneficial way.

Taking the Fourier transform, we let u_k be the Fourier coefficients of u. We get the evolution equations

$$\partial_t u_k = ik^5 u_k + i\omega k u_k - ik \sum_{j=-\infty}^{\infty}{}' u_{k-j} u_j.$$

We only consider $k \neq 0$ since we know already that $\partial_t u_0 = u_0 = 0$. The prime on the sum indicates that $j = 0$ and $j = k$ are excluded, as these modes are unnecessary since the mean of u is equal to zero. The initial condition is

$$u_k(0) = \hat{g}(k).$$

We bring the linear terms to the left-hand side:

$$\partial_t u_k - ik^5 u_k - i\omega k u_k = -ik \sum_{j=-\infty}^{\infty}{}' u_{k-j} u_j.$$

We use an integrating factor, so we define v_k through the equation

$$v_k(t) = u_k(t) e^{-ik^5 t - i\omega k t}.$$

This yields

$$\partial_t v_k = -ik \sum_{j=-\infty}^{\infty}{}' e^{-ik^5 t - i\omega k t} e^{i(k-j)^5 t + i\omega(k-j)t} e^{ij^5 t + i\omega j t} v_{k-j} v_j.$$

The exponents simplify as the terms in the exponent related to the transport speed ω all cancel. This leaves us with

$$\partial_t v_k = -ik \sum_{j=-\infty}^{\infty}{}' e^{-ik^5 t} e^{i(k-j)^5 t} e^{ij^5 t} v_{k-j} v_j.$$

It will therefore be helpful to understand the quantity $k^5 - (k-j)^5 - j^5$, as this appears in the exponent. Elementary calculations show that, upon introducing the notation $\sigma = k^2 - jk + j^2$, we have

$$k^5 - (k-j)^5 - j^5 = 5(k-j)jk\sigma.$$

Using this identity, we are able to write the evolution equation for the v_k,

$$\partial_t v_k = -ik \sum_{j=-\infty}^{\infty}{}' e^{-5i(k-j)jk\sigma t} v_{k-j} v_j. \tag{21}$$

We can then rewrite this, recognizing that the exponential is in fact the derivative of an exponential:

$$\partial_t v_k = -ik \sum_{j=-\infty}^{\infty}{}' \left(\partial_t \left(\frac{e^{-5i(k-j)jk\sigma t}}{-5i(k-j)jk\sigma}\right)\right) v_{k-j} v_j = \sum_{j=-\infty}^{\infty}{}' \left(\partial_t \left(\frac{e^{-5i(k-j)jk\sigma t}}{5(k-j)j\sigma}\right)\right) v_{k-j} v_j.$$

Next, we "differentiate by parts", moving the time derivative:

$$\partial_t v_k = \partial_t\left(\frac{1}{5}\sum_{j=-\infty}^{\infty}{}' \frac{e^{-5i(k-j)jk\sigma t}}{(k-j)j\sigma} v_{k-j}v_j\right) - \frac{1}{5}\sum_{j=-\infty}^{\infty}{}' \frac{e^{-5i(k-j)jk\sigma t}}{(k-j)j\sigma}\partial_t(v_{k-j}v_j).$$

We define B through its Fourier coefficients, $B_k(t)$, as

$$B_k(t) = -\frac{1}{5}\sum_{j=-\infty}^{\infty}{}' \frac{e^{-5i(k-j)jk\sigma t}}{(k-j)j\sigma} v_{k-j}v_j. \tag{22}$$

We then are able to write the evolution equations as

$$\partial_t[v_k + B_k] = -\frac{1}{5}\sum_{j=-\infty}^{\infty}{}' \left[\frac{e^{-5i(k-j)jk\sigma t}}{(k-j)j\sigma}\right]\left[(\partial_t v_{k-j})v_j + v_{k-j}(\partial_t v_j)\right].$$

Next, we substitute for $\partial_t v_{k-j}$ and $\partial_t v_j$. We let $\tilde{\sigma} = j^2 - j\ell + \ell^2$; using (21), we have

$$\partial_t v_j = -ij\sum_{\ell=-\infty}^{\infty}{}' e^{-5itj\ell(j-\ell)\tilde{\sigma}} v_{j-\ell}v_\ell.$$

We are then able, using a symmetry between $k-j$ and j, to write

$$\partial_t[v_k + B_k] = \frac{2i}{5}\sum_{j=-\infty}^{\infty}{}' \sum_{\ell=-\infty}^{\infty}{}' \left[\frac{e^{-5it[(k-j)jk\sigma + (j-\ell)j\ell\tilde{\sigma}]}}{(k-j)\sigma}\right] v_{k-j}v_{j-\ell}v_\ell. \tag{23}$$

(We reiterate that $k \neq j$.) We give the name R_k to the right-hand side of (23), and we let R be the function with Fourier coefficients equal to R_k for all k. So, we have $\partial_t[v_k + B_k] = R_k$. Integrating with respect to time, we have

$$v_k(t) - v_k(0) = B_k(0) - B_k(t) + \int_0^t R_k(s)\,ds. \tag{24}$$

We transform back to u by multiplying (24) by $e^{ik^5t + i\omega kt}$, and we note that $v_k(0) = u_k(0)$. These considerations yield

$$u_k(t) - e^{ik^5t+i\omega kt}u_k(0) = e^{ik^5t+i\omega kt}\left(B_k(0) - B_k(t) + \int_0^t R_k(s)\,ds\right). \tag{25}$$

Notice that (25) is the k-th Fourier coefficient of the Duhamel integral at time t, or $\mathcal{F}(S_D(t)u_0)(k)$.

3.3. *Estimates.* We now estimate B and R and associated quantities, to demonstrate the smoothing described in Theorem 4 for our equation with fifth-order dispersion.

Remark 8. It will be plain to see that all estimates made in the present section are uniform in ω.

Lemma 9. *If $s \geq 1$ and $v \in H^s$, then $B \in H^{s+3}$, with the estimate*

$$\|B\|_{H^{s+3}}^2 \leq c\|v\|_{H^s}^4.$$

(We note that for our main theorem, we only actually need $B \in H^{s+2}$, but it turns out that $B \in H^{s+3}$.)

Proof. We will show that $\partial_x^3 B$ is in H^s. We begin by taking three derivatives of B, which requires multiplying (22) by $(ik)^3$:

$$(ik)^3 B_k = \frac{i}{5} \sum_{j=-\infty}^{\infty}{}' E(k,j) \frac{k^3}{j(k-j)(k^2-kj+j^2)} v_{k-j} v_j,$$

where $E(k,j)$ represents the exponential, $E(k,j) = e^{-5itjk(k-j)\sigma}$.

We will demonstrate now that

$$\frac{k^3}{j(k-j)(k^2-kj+j^2)}$$

is bounded by a constant. To begin, we consider

$$\left|\frac{k}{j(k-j)}\right| = \left|\frac{k-j+j}{j(k-j)}\right| \le \left|\frac{1}{j}\right| + \left|\frac{1}{k-j}\right| \le 2.$$

Next, we consider $\sigma = k^2 - kj + j^2$. We observe that

$$\sigma = \tfrac{1}{2}k^2 + \tfrac{1}{2}(k-j)^2 + \tfrac{1}{2}j^2, \tag{26}$$

so clearly $\sigma \ge \frac{1}{2}k^2$. Thus,

$$\left|\frac{k^2}{\sigma}\right| \le 2. \tag{27}$$

Thus, for any k and any j, we have

$$\left|E(k,j)\frac{k^3}{j(k-j)(k^2-kj+j^2)}\right| \le 4. \tag{28}$$

We give the name

$$\Phi(k,j) = E(k,j)\frac{k^3}{j(k-j)(k^2-kj+j^2)}.$$

Of course, we have

$$\|\partial_x^{s+3} B\|_{L^2}^2 = \big\||k|^s (ik^3 B_k)\big\|_{\ell^2}^2.$$

We then have

$$\big\||k|^s (ik^3 B_k)\big\|_{\ell^2}^2 = \frac{1}{25} \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty}{}' \sum_{\ell=-\infty}^{\infty}{}' \Phi(k,j)\bar{\Phi}(k,\ell) k^{2s} v_{k-j} v_j \bar{v}_{k-\ell} \bar{v}_\ell.$$

In light of (28), we have

$$\big\||k|^s (ik^3 B_k)\big\|_{\ell^2}^2 \le \sum_{k,j,\ell} k^{2s} |v_{k-j}||v_j||v_{k-\ell}||v_\ell|. \tag{29}$$

Let V be the function defined through its Fourier coefficients as $\mathcal{F}V(k) = |v_k|$ for all k. Note that since $v \in H^s$, we have $V \in H^s$, with $\|v\|_{H^s} = \|V\|_{H^s}$. Since $V \in H^s$ with $s \ge 1$, we can see that $V^2 \in H^s$, with $\|V^2\|_{H^s} \le c\|v\|_{H^s}^2$. Notice that the right-hand side of (29) is equal to $\|\partial_x^s (V^2)\|_{L^2}^2$. This completes the proof. □

The particular estimate we need for B follows from Lemma 9:

Corollary 10. *If $s \geq 1$ and $u \in C([0,T]; H^s)$, and if $t \in [0,T]$, then*

$$\mathcal{F}^{-1}\big(e^{ik^5t+i\omega kt}(B_k(0) - B_k(t))\big) \in H^{s+3},$$

with the bound

$$\big\|\mathcal{F}^{-1}\big(e^{ik^5t+i\omega kt}(B_k(0) - B_k(t))\big)\big\|_{H^{s+3}} \leq c\|u\|^2_{C([0,T];H^s)}.$$

Proof. This follows immediately from Lemma 9, and from the fact that $|e^{i\theta}| = 1$ for any real θ. ☐

Having proved a satisfactory estimate for B, we turn to R.

Lemma 11. *If $s \geq 1$ and $v \in H^s$, then $R \in H^{s+2}$, with the estimate*

$$\|R\|^2_{H^{s+2}} \leq c\|v\|^6_{H^s}.$$

Proof. Recall the formula for R_k,

$$R_k = \frac{2i}{5}\sum_{j=-\infty}^{\infty}{}'\sum_{\ell=-\infty}^{\infty}{}'\frac{1}{(k-j)\sigma}(v_{k-j}v_{j-\ell}v_\ell)\exp\big\{(-5it)(kj(k-j)\sigma + j\ell(j-\ell)\tilde{\sigma})\big\},$$

with $\sigma = k^2 - kj + j^2$ and $\tilde{\sigma} = j^2 - j\ell + \ell^2$. As we noted in the proof of Lemma 9, we have $k^2/\sigma \leq 2$. Following the lines of the proof of Lemma 9, we arrive at

$$\|\partial_x^{s+2}R\|^2_{L^2} \leq c\sum_{k,j,\ell,m,n} k^{2s}|v_{k-j}||v_{j-\ell}||v_\ell||v_{k-m}||v_{m-n}||v_n|.$$

Letting V be as in the proof of Lemma 9, we see that the right-hand side is a constant times the square of the L^2 norm of $\partial_x^s(V^3)$. Thus, this is bounded by $\|v\|^6_{H^s}$, as claimed. ☐

Lemma 11 implies the following, which is the estimate we need for R:

Corollary 12. *If $s \geq 1$ and $u \in C([0,T]; H^s)$, and if $t \in [0,T]$, then*

$$\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right) \in H^{s+2},$$

with

$$\left\|\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right)\right\|_{H^{s+2}} \leq c\|u\|^3_{C([0,T];H^s)}.$$

Proof. We begin by noting that, of course,

$$\begin{aligned}\left\|\partial_x^{s+2}\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right)\right\|^2_{L^2} &= \left\|(ik)^{s+2}e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right\|^2_{\ell^2}\\ &= \left\|\int_0^t k^{s+2}R_k(s)\,ds\right\|^2_{\ell^2}\\ &= \sum_{k=-\infty}^{\infty}\left(\int_0^t k^{s+2}R_k(s)\,ds\right)\left(\int_0^t k^{s+2}\bar{R}_k(\tau)\,d\tau\right).\end{aligned}$$

We use the triangle inequality:

$$\left\|\partial_x^{s+2}\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right)\right\|_{L^2}^2 \leq \sum_{k=-\infty}^{\infty}\left(\int_0^t |k^{s+2}R_k(s)|\,ds\right)\left(\int_0^t |k^{s+2}R_k(\tau)|\,d\tau\right).$$

By Tonelli's theorem, we can exchange the sum and the integrals:

$$\left\|\partial_x^{s+2}\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right)\right\|_{L^2}^2 \leq \int_0^t\int_0^t \sum_{k=-\infty}^{\infty}|k^{s+2}R_k(s)||k^{s+2}R_k(\tau)|\,ds\,d\tau.$$

We then use the Cauchy–Schwarz inequality:

$$\left\|\partial_x^{s+2}\mathcal{F}^{-1}\left(e^{ik^5t+i\omega kt}\int_0^t R_k(s)\,ds\right)\right\|_{L^2}^2 \leq c\int_0^t\int_0^t \|R(\,\cdot\,,s)\|_{H^{s+2}}\|R(\,\cdot\,,\tau)\|_{H^{s+2}}\,ds\,d\tau.$$

We then use Lemma 11, and the proof is complete. □

We are now in a position to show that the necessary estimate for the Duhamel integral holds for (17).

Theorem 13. *Let $0 < T_1 < T_2$ be given. There exists $\gamma > 0$ such that for any $u_0 \in H^6$ such that $\|u_0\|_{H^6} < \gamma$, there is a unique solution of the initial value problem* (17) *with initial data u_0, with the solution $u \in C([0,T_2]; H^6)$. There exists $c > 0$ and $\tilde{\gamma} \in (0,\gamma)$ such that for any $T \in [T_1, T_2]$, and for any $u_0 \in H_0^6$ such that $\|u_0\|_{H^6} < \tilde{\gamma}$, we have*

$$\|S_D(T)u_0\|_{H^8} \leq c\|u_0\|_{H^6}^2.$$

Proof. The formula (25) and the estimates of Corollaries 10 and 12 for B and R immediately imply

$$\|S_D(T)u_0\|_{H^8} \leq c\|u\|_{C([0,T];H^6)}^2.$$

However, we are not yet finished because we need the bound to be in terms of the initial data, and not in terms of the solution at positive times.

As discussed in Section 3.1 above, we have

$$\frac{d}{dt}\|u^\varepsilon\|_{H^6} \leq c\|u^\varepsilon\|_{H^6}^2.$$

Let $\|u_0\|_{H^6} = \delta/2$. Then, as long as $\|u^\varepsilon(\,\cdot\,,t)\|_{H_6} \leq \delta = 2\|u_0\|_{H^6}$, we have

$$\frac{d}{dt}\|u^\varepsilon\|_{H^6} \leq c\delta\|u^\varepsilon\|_{H^6},$$

and thus

$$\|u^\varepsilon\|_{H^6} \leq \|u_0\|_{H^6}e^{c\delta t} = \tfrac{1}{2}\delta e^{c\delta t}.$$

This implies that $\|u^\varepsilon\|_{H^6} \leq 2\|u_0\|_{H^6}$ as long as $e^{c\delta t} < 2$. This is valid as long as $t < \ln(2)/c\delta$; notice that this bound goes to infinity as δ vanishes (that is, the "doubling time" for solutions goes to infinity as the initial size of the solutions goes to zero). Taking the limit as ε vanishes, then (along our subsequence), we find

$$\|u\|_{H^6} \leq 2\|u_0\|_{H^6},$$

as long as $t < \ln(2)/c\delta$.

Given the set of potential temporal periods of interest, $[T_1, T_2]$, we may choose δ sufficiently small so that as long as $\|u_0\|_{H^6} < \delta/2$, for all $t \in [0, T_2]$, we have $\|u(\cdot, t)\|_{H^6} < 2\|u_0\|_{H^6}$. We have shown above that $\|S_D(T)u_0\|_{H^8} \leq c\|u\|^2_{C([0,T_2];H^6)}$. We now can bound this in terms of $\|u_0\|_{H^6}$, so that

$$\|S_D(T)u_0\|_{H^8} \leq 4c\|u_0\|^2_{H^6}$$

for any $T \in [T_1, T_2]$ and for any u_0 satisfying our smallness assumption. This is the desired bound. ☐

3.4. ***Completion of the example.*** In this section, we state a specific theorem on nonexistence of time-periodic solutions, making use of the above. Consider the equation

$$\partial_t \tilde{u} = \partial_x^5 \tilde{u} - 2\tilde{u}\, \partial_x \tilde{u}. \tag{30}$$

Consider the initial data, $\tilde{u}_0 \in H^6$, with the mean of $\tilde{u}_0$ equal to α. As above, we define $u_0 = \tilde{u}_0 - \alpha$, and we let $u = \tilde{u} - \alpha$. The evolution equation satisfied by u, as discussed above, is

$$\partial_t u = \partial_x^5 u - 2\alpha\, \partial_x u - 2u\, \partial_x u. \tag{31}$$

If we let $A = \partial_x^5 - 2\alpha\partial_x$, then we see that the symbol of A is

$$\hat{A}(k) = i(k^5 - 2\alpha k) = ik(k^4 - 2\alpha). \tag{32}$$

Thus, if $2\alpha < 1$, then there are no zeros of the symbol in $\mathbb{Z} \setminus \{0\}$. Recalling the hypotheses **(H)**, we have $M = 2$, and we let $r_1 = 5$, $r_2 = 1$, $Z_1 = \{1\}$, and $Z_2 = \left(-\frac{1}{2}, \frac{1}{2}\right)$. Letting $\alpha_1 \in Z_1$ and $\alpha_2 \in Z_2$, we see that we may take $\beta_1 = \frac{1}{2}$ and $\beta_2 = \frac{1}{2}$. Letting $\alpha_2 = -2\alpha$, we see that the hypotheses **(H)** are satisfied, and Lemma 2 and Corollary 3 hold, with uniform estimates for $\alpha \in \left(-\frac{1}{4}, \frac{1}{4}\right)$. Therefore, Theorem 4 applies for $\alpha \in \left(-\frac{1}{4}, \frac{1}{4}\right)$. Theorem 13 also applies, and in light of Remarks 7 and 8, we see that the constants in Theorem 13 can be taken to be uniform with respect to α. This implies that (31) does not possess small, nonzero time-periodic solutions, uniformly in $\alpha \in \left(-\frac{1}{4}, \frac{1}{4}\right)$.

Adding the mean α back to u, we get $\tilde{u} = (\tilde{u} - \alpha) + \alpha$, with $u = \tilde{u} - \alpha$. We know that for $\alpha \in \left(-\frac{1}{4}, \frac{1}{4}\right)$, $\tilde{u} - \alpha$ does not possess small, nonzero time-periodic solutions with the associated periods ($T \in W_{p,\delta}$ or $T \in W_p$, as appropriate). Furthermore, we know that $\|\tilde{u}\|_{H^6} \geq \|\tilde{u} - \alpha\|_{H^6}$. This implies that the only small time-periodic solutions of (30) with the given temporal periods are $\tilde{u} = \alpha$. Thus (30) does not possess small, nonconstant time-periodic solutions with the given temporal periods. This proves the following corollary:

Corollary 14. *Let* $0 < T_1 < T_2$ *be given. Let* $p \in (1, 2]$ *be given. Let* $0 < \delta < T_2 - T_1$ *be given. Let* $W_{p,\delta} \subseteq [T_1, T_2]$ *be as in Lemma 2, with* A *given by* (32) *for* $|\alpha| \leq \frac{1}{4}$. *There exists* $r_1 > 0$ *such that for all* $T \in W_{p,\delta}$, *if* u *is a smooth, nonconstant time-periodic solution of* (30) *with temporal period* T, *then*

$$\inf_{t \in [0,T]} \|u\|_{H^6} > r_1.$$

Let $W_p \subseteq [T_1, T_2]$ *be as in Corollary 3. Let* $T \in W_p$ *be given. Then there exists* r_2 *such that if* u *is a smooth, nonconstant time-periodic solution of* (30) *with temporal period* T, *then*

$$\inf_{t \in [0,T]} \|u\|_{H^6} > r_2.$$

4. Further examples with $\tilde{p} = 0$

In this section, we provide a few other equations which can be treated similarly to the above. We do not provide full proofs in this section, but instead point out the differences with the prior proof.

4.1. ***Nonresonant Kawahara equations.*** The Kawahara equation has been justified as a model for water waves with surface tension [Düll 2012; Schneider and Wayne 2002]. It can be written as

$$\partial_t \tilde{u} = \partial_x^5 \tilde{u} - \theta\, \partial_x^3 \tilde{u} - 2\tilde{u}\, \partial_x \tilde{u}, \tag{33}$$

with $\theta > 0$. As before, we take this with initial data $\tilde{u}(\cdot, 0) = \tilde{u}_0$, and we assume that the mean of $\tilde{u}_0$ is equal to α. We again let $u = \tilde{u} - \alpha$, and we find that the equation satisfied by u is

$$\partial_t u = \partial_x^5 u - \theta\, \partial_x^3 u - 2\alpha\, \partial_x u - 2u\, \partial_x u. \tag{34}$$

Our prior results extend to the Kawahara equation as long as the constant θ is chosen to avoid resonance. In particular, we must require $k^5 - \theta k^3 \neq 0$ for all $k \in \mathbb{Z} \setminus \{0\}$. Notice that this is the same as requiring

$$\min_{k \in \mathbb{Z} \setminus \{0\}} |k^5 - \theta k^3| > 0. \tag{35}$$

This implies that there exists a constant $\bar{\alpha} > 0$ and a constant $\beta_2 > 0$ such that for all $\alpha \in (-\bar{\alpha}, \bar{\alpha})$, for all $k \in \mathbb{Z} \setminus \{0\}$,

$$|k^5 - \theta k^3 - 2\alpha k| \geq \beta_2.$$

We now verify **(H)**, taking $M = 3$. We let $r_1 = 5$, $r_2 = 3$, and $r_3 = 1$. We let $Z_1 = \{1\}$ and $Z_2 = \{\theta\}$. We take $\alpha_1 = 1$, $\alpha_2 = \theta$, and $\alpha_3 = -2\alpha$, with $Z_3 = (-2\bar{\alpha}, 2\bar{\alpha})$. Then, **(H)** is satisfied, with $\beta_1 = \frac{1}{2}$.

This means that Theorem 4 applies, and all that must be done to conclude the nonexistence of small doubly periodic waves for the nonresonant Kawahara equation is that the smoothing property must be demonstrated. We are able to demonstrate the smoothing property with $\tilde{p} = 0$.

We take the Fourier transform of (34), finding

$$\partial_t u_k = ik^5 u_k + i\theta k^3 u_k - 2i\alpha k u_k - ik \sum_{j=-\infty}^{\infty} u_{k-j} u_j.$$

As before, we use an integrating factor, defining

$$v_k = u_k \exp\bigl\{it(-k^5 - \theta k^3 + 2\alpha k)\bigr\}.$$

We have the following evolution equation for v_k:

$$\partial_t v_k = -ik {\sum_{j=-\infty}^{\infty}}' (v_{k-j} v_j) \exp\{it\Phi(j,k)\},$$

where the phase function, Φ, is given by

$$\Phi(j,k) = -k^5 - \theta k^3 + 2\alpha k + (k-j)^5 + \theta(k-j)^3 - 2\alpha(k-j) + j^5 + \theta j^3 - 2\alpha j.$$

This simplifies, as all the terms with α cancel, and also since we previously computed $k^5-(k-j)^5-j^5$. We have not previously computed $k^3-(k-j)^3-j^3$, but this is straightforward:

$$k^3-(k-j)^3-j^3=3jk(k-j).$$

These considerations imply

$$\Phi(j,k)=5jk(k-j)\sigma+3\theta jk(k-j)=5jk(k-j)\big(\sigma+\tfrac{3}{5}\theta\big).$$

The critical step in the proof of smoothing in Section 3 was inequality (27). We see that the corresponding inequality in the present case is

$$\left|\frac{k^2}{\sigma+\frac{3}{5}\theta}\right|\leq 2,$$

which follows immediately from (26) and the condition $\theta>0$. The rest of the proof of Section 3 can be repeated, establishing the following:

Corollary 15. *Let $0<T_1<T_2$ be given. Let $p\in(1,2]$ be given. Let $0<\delta<T_2-T_1$ be given. Let $\theta>0$ satisfy* (35), *and let $\bar{\alpha}>0$ be as above. Let $W_{p,\delta}\subseteq[T_1,T_2]$ be as in Lemma 2, with A given by $\mathcal{F}(A)=i(-k^5-\theta k^3+2\alpha k)$ for $|\alpha|\leq\bar{\alpha}$. There exists $r_1>0$ such that for all $T\in W_{p,\delta}$, if u is a smooth, nonconstant time-periodic solution of* (33) *with temporal period T, then*

$$\inf_{t\in[0,T]}\|u\|_{H^6}>r_1.$$

Let $W_p\subseteq[T_1,T_2]$ be as in Corollary 3. Let $T\in W_p$ be given. Then there exists r_2 such that if u is a smooth, nonconstant time-periodic solution of (33) *with temporal period T, then*

$$\inf_{t\in[0,T]}\|u\|_{H^6}>r_2.$$

Remark 16. Above, we appeared to use in a fundamental way the property $\theta>0$. We assume $\theta>0$ only because this appears to be a feature of the Kawahara equation as it exists in the prior literature. If we instead had $\theta<0$, our argument would still work. For a particular value of θ, if there exist $(j,k)\in(\mathbb{Z}\setminus\{0\})^2$ such that $\sigma+\frac{3}{5}\theta=0$, then we treat such values of (j,k) differently. The arguments of Section 3 continue to apply whenever this quantity does not vanish. All that remains is to observe that, as can be seen from (26), the set of (j,k) for which $\sigma+\frac{3}{5}\theta$ does vanish is bounded for a fixed value of θ (or for values of θ in a bounded set), and that regularity is determined by behavior for large k.

4.2. ***Seventh-order equations.*** The proof of smoothing for the fifth-order equation above depended on the factorization of $k^5-(k-j)^5-j^5$. With dispersion of seventh order, we must understand $k^7-(k-j)^7-j^7$. By making elementary calculations, one may show that

$$k^7-(k-j)^7-j^7=7(k-j)jk\tau, \tag{36}$$

where

$$\tau=\tfrac{1}{2}k^4+\tfrac{1}{2}(k-j)^4+\tfrac{1}{2}j^4;$$

note the similarity to (26). Clearly, we have the inequality

$$\left|\frac{k^4}{\tau}\right| = \frac{k^4}{\tau} \le 2. \tag{37}$$

We can then perform all of the calculations of Section 3 for the equation

$$\partial_t u = \partial_x^7 u - 2u\,\partial_x u. \tag{38}$$

Recall that B and R, defined in Section 3.2, were shown in Section 3.3 to gain two derivatives. This property hinged on the inequality (27). In the present setting, the analogues of B and R would now gain four derivatives because of (37). (As noted in the introduction, this allows one to see that the smoothing mechanism we are using is different than Kato smoothing, as Kato smoothing would provide a gain of three derivatives with seventh-order dispersion.) Following the argument of Section 3, but using (37), we arrive at the following:

Corollary 17. *Let* $0 < T_1 < T_2$ *be given. Let* $p \in (1, 4]$ *be given. Let* $0 < \delta < T_2 - T_1$ *be given. Let* $W_{p,\delta} \subseteq [T_1, T_2]$ *be as in Lemma 2, with* A *given by* $\mathcal{F}(A) = i(-k^7 + \alpha k)$ *for* $|\alpha| \le \frac{1}{2}$. *There exists* $r_1 > 0$ *such that for all* $T \in W_{p,\delta}$, *if* u *is a smooth, nonconstant time-periodic solution of* (38) *with temporal period* T, *then*

$$\inf_{t\in[0,T]} \|u\|_{H^8} > r_1.$$

Let $W_p \subseteq [T_1, T_2]$ *be as in Corollary 3. Let* $T \in W_p$ *be given. Then there exists* r_2 *such that if* u *is a smooth, nonconstant time-periodic solution of* (38) *with temporal period* T, *then*

$$\inf_{t\in[0,T]} \|u\|_{H^8} > r_2.$$

We note that a change from Corollary 17 as compared to Corollary 14 is that we are now using H^8 instead of H^6. The function space is chosen so that the solutions under consideration are classical solutions; using the space H^8, the seventh derivative of u appearing on the right-hand side of (38) is classically defined.

5. The KdV equation

We study the equation

$$\partial_t \tilde{u} = -\partial_x^3 \tilde{u} - \tilde{u}\,\partial_x \tilde{u}. \tag{39}$$

(This is not the most traditional choice of coefficients for the KdV equation, but the coefficients are changeable by scaling, so it makes no difference.) The evolution equation (39) is taken with initial data $u(\cdot, 0) = \tilde{g}$. We let the mean of $\tilde{g}$ be denoted as α, and we define $g = \tilde{g} - \alpha$. Then, noticing that (39) preserves the mean of the solution, we define $u = \tilde{u} - \alpha$, so that the mean of u is equal to zero, as long as the solution $\tilde{u}$ of (39) exists. The evolution equation satisfied by u is

$$\partial_t u = -\partial_x^3 u - \alpha\,\partial_x u - u\,\partial_x u. \tag{40}$$

We note that there are very many papers in the literature treating the well-posedness of the KdV equation. For example, in [Bourgain 1993], global well-posedness for the periodic KdV equation with initial data in H^s for $s \geq 0$ is established, and in [Kappeler and Topalov 2006], global well-posedness is established in H^s for $s \geq -1$. Nevertheless, we remark that the simple well-posedness proof for (16) given in Section 3.1 can be suitably and straightforwardly modified to the KdV equation, yielding the same results: that classical solutions exist for arbitrarily long intervals of time if the initial data is sufficiently small, and that the doubling time of such solutions goes to infinity as the size of the data vanishes.

We take the Fourier transform of (40):

$$\partial_t u_k = i(k^3 - \alpha k)u_k - i \sum_{j=-\infty}^{\infty}{}' u_{k-j}(ju_j).$$

Since the mean of u is equal to zero, the values $j = 0$ and $j = k$ are unnecessary; as before, the prime indicates that these indices are excluded from the summation. We use an integrating factor, defining $v_k = u_k \exp\{-it(k^3 - \alpha k)\}$:

$$\partial_t v_k = -i \sum_{j=-\infty}^{\infty}{}' \exp\big\{-it(k^3 - (k-j)^3 - j^3)\big\} v_{k-j}(jv_j).$$

(Notice that the terms in the exponential involving α all canceled.) From Pascal's triangle, we see that $k^3 - (k-j)^3 - j^3 = 3kj(k-j)$. Thus, we may write

$$\partial_t v_k = -i \sum_{j=-\infty}^{\infty}{}' \exp\big\{-3itkj(k-j)\big\} v_{k-j}(jv_j). \tag{41}$$

We now manipulate the exponential:

$$\partial_t v_k = -i \sum_{j=-\infty}^{\infty}{}' \partial_t \left(\frac{e^{-3itkj(k-j)}}{-3ikj(k-j)} \right) v_{k-j}(jv_j).$$

We cancel the factor of $-i$, and we also "differentiate by parts":

$$\partial_t v_k = \partial_t \left(\sum_{j=-\infty}^{\infty}{}' \frac{e^{-3itkj(k-j)}}{3kj(k-j)} v_{k-j}(jv_j) \right) - \sum_{j=-\infty}^{\infty} \frac{e^{-3itkj(k-j)}}{3kj(k-j)} \partial_t (v_{k-j}(jv_j)).$$

We can write this as

$$\partial_t (v_k + B_k) = R_k,$$

with

$$B_k = -\frac{1}{3} \sum_{j=-\infty}^{\infty} \frac{e^{-3itkj(k-j)}}{kj(k-j)} v_{k-j}(jv_j),$$

$$R_k = -\frac{1}{3} \sum_{j=-\infty}^{\infty}{}' \frac{e^{-3itkj(k-j)}}{kj(k-j)} \partial_t (v_{k-j}(jv_j)). \tag{42}$$

We continue to rewrite this, by applying the time derivative on the right-hand side of (42). Using (41), we have

$$\partial_t v_{k-j} = -i \sum_{m=-\infty}^{\infty}{}' \exp\{-3it(k-j)m(k-j-m)\} v_{k-j-m}(mv_m),$$
$$\partial_t v_j = -i \sum_{\ell=-\infty}^{\infty}{}' \exp\{-3itj\ell(j-\ell)\} v_{j-\ell}(\ell v_\ell).$$

Using these, we write $R_k = R_k^1 + R_k^2$, with

$$R_k^1 = \frac{i}{3} \sum_{j=-\infty}^{\infty}{}' \sum_{m=-\infty}^{\infty}{}' \frac{\exp\{-3it(k-j)(kj+m(k-j-m))\}}{kj(k-j)} v_{k-j-m}(jv_j)(mv_m),$$
$$R_k^2 = \frac{i}{3} \sum_{j=-\infty}^{\infty}{}' \sum_{\ell=-\infty}^{\infty}{}' \frac{\exp\{-3itj(k(k-j)+\ell(j-\ell))\}}{kj(k-j)} v_{k-j}(j(v_{j-\ell})(\ell v_\ell)).$$

We further rewrite R_k^2 by writing $j = j - \ell + \ell$:

$$R_2^k = \frac{i}{3} \sum_{j=-\infty}^{\infty}{}' \sum_{\ell=-\infty}^{\infty}{}' \frac{\exp\{-3itj(k(k-j)+\ell(j-\ell))\}}{kj(k-j)} v_{k-j}((j-\ell)v_{j-\ell})(\ell v_\ell)$$
$$+ \frac{i}{3} \sum_{j=-\infty}^{\infty}{}' \sum_{\ell=-\infty}^{\infty}{}' \frac{\exp\{-3itj(k(k-j)+\ell(j-\ell))\}}{kj(k-j)} v_{k-j} v_{j-\ell}(\ell^2 v_\ell).$$

Using the inequality

$$\left| \frac{k^2}{kj(k-j)} \right| \le 2,$$

and our previous arguments, we find the bounds

$$\|B\|_{H^{s+2}} \le c\|v\|_{H^s}\|v\|_{H^{s+1}},$$
$$\|R^1\|_{H^{s+2}} \le c\|v\|_{H^s}\|v\|_{H^{s+1}}^2,$$
$$\|R^2\|_{H^{s+2}} \le c\|v\|_{H^s}\big(\|v\|_{H^{s+1}}^2 + \|v\|_{H^s}\|v\|_{H^{s+2}}\big).$$

If $\|v\|_{H^{s+2}} < 1$, then the right-hand sides can all be bounded by

$$c\|v\|_{H^s}\|v\|_{H^{s+2}}.$$

Following the arguments of the previous cases, we are able to conclude, for sufficiently small $u_0 \in H_0^{s+2}$,

$$\|S_D(T)u_0\|_{H^{s+2}} \le c\|u_0\|_{H^s}\|u_0\|_{H^{s+2}}.$$

This is estimate (14) with $p = \tilde{p} = 2$ and $q = 1$, so we see that the nonexistence result holds for the KdV equation:

Corollary 18. *Let $0 < T_1 < T_2$ be given. Let $p \in (1,2]$ be given. Let $0 < \delta < T_2 - T_1$ be given. Let $W_{p,\delta} \subseteq [T_1, T_2]$ be as in Lemma 2, with A given by*

$$A = -\partial_x^3 - \alpha\,\partial_x,$$

with $|\alpha| \le \frac{1}{2}$. There exists $r_1 > 0$ such that for all $T \in W_{p,\delta}$, if u is a smooth, nonconstant time-periodic solution of (39) *with temporal period T, then*

$$\inf_{t\in[0,T]} \|u\|_{H^6} > r_1.$$

Let $W_p \subseteq [T_1, T_2]$ be as in Corollary 3. Let $T \in W_p$ be given. Then there exists r_2 such that if u is a smooth, nonconstant time-periodic solution of (39) *with temporal period T, then*

$$\inf_{t\in[0,T]} \|u\|_{H^6} > r_2.$$

We note that the result of Corollary 18 is given for solutions in H^6. This is because we take $s = 4$ so that we work with classical solutions, and we have $\tilde{p} = 2$. An application of Theorem 4 then gives a restriction on the $H^{s+\tilde{p}} = H^6$ norm.

6. Discussion

We have developed a theoretical framework for the demonstration of nonexistence of small doubly periodic solutions for dispersive evolution equations. The abstract theorem indicates that nonexistence follows from the demonstration of dispersive smoothing estimates. In particular cases, we have demonstrated that the required dispersive smoothing estimates hold. These results are an analogue of scattering results for dispersive equations on the real line, since scattering implies the nonexistence of small-amplitude coherent structures.

Other work to be done includes treating additional specific examples, and possibly proving a general theorem about when the weak smoothing property (14) holds. For example, it should be investigated whether the necessary properties can be shown to hold for other dispersive equations in one space dimension (like the Benjamin–Ono equation) and in higher dimensions (such as Schrödinger equations). For equations with sufficiently strong dispersion, the stronger smoothing estimate (corresponding to $\tilde{p} = 0$, in which the Duhamel integral gains more than one derivative) will likely hold. A class of equations with strong dispersion are fourth-order Schrödinger equations (see, e.g., [Fibich et al. 2002; 2003; Karpman and Shagalov 2000]). Such equations are of the form

$$i\psi_t + \Delta\psi + |\psi|^{2\sigma}\psi + \varepsilon\Delta^2\psi = 0,$$

with $\sigma > 0$ and $\varepsilon > 0$, and can arise by including higher-order corrections when deriving a Schrödinger equation from the Maxwell equations. The above linear estimates, such as Lemma 2 and Corollary 3, are valid in one spatial dimension. As pointed out to the authors by the referee of [Ambrose and Wright 2014], in n spatial dimensions, the result requires $p > n$ rather than $p > 1$. Thus, in higher dimensions, one would expect to need to use the weaker smoothing property (corresponding to $\tilde{p} > 0$) in order to follow the strategy of the present paper.

Acknowledgment

We thank the referee for a careful reading of the manuscript and for helpful remarks.

References

[Alazard and Baldi 2015] T. Alazard and P. Baldi, "Gravity capillary standing water waves", *Arch. Ration. Mech. Anal.* **217**:3 (2015), 741–830. MR 3356988 Zbl 1317.35181

[Ambrose and Wilkening 2009] D. M. Ambrose and J. Wilkening, "Global paths of time-periodic solutions of the Benjamin–Ono equation connecting pairs of traveling waves", *Commun. Appl. Math. Comput. Sci.* **4** (2009), 177–215. MR 2010j:35437 Zbl 1184.35271

[Ambrose and Wilkening 2010] D. M. Ambrose and J. Wilkening, "Computation of time-periodic solutions of the Benjamin–Ono equation", *J. Nonlinear Sci.* **20**:3 (2010), 277–308. MR 2011d:65146 Zbl 1203.37085

[Ambrose and Wright 2014] D. M. Ambrose and J. D. Wright, "Non-existence of small-amplitude doubly periodic waves for dispersive equations", *C. R. Math. Acad. Sci. Paris* **352**:7-8 (2014), 597–602. MR 3237811 Zbl 1301.35128

[Baldi 2013] P. Baldi, "Periodic solutions of fully nonlinear autonomous equations of Benjamin–Ono type", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **30**:1 (2013), 33–77. MR 3011291 Zbl 1285.35090

[Bourgain 1993] J. Bourgain, "Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, II: The KdV-equation", *Geom. Funct. Anal.* **3**:3 (1993), 209–262. MR 95d:35160b Zbl 0787.35098

[Bourgain 1994] J. Bourgain, "Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE", *Internat. Math. Res. Notices* **1994**:11 (1994), 475–497. MR 96f:58170 Zbl 0817.35102

[Bourgain 1995a] J. Bourgain, "Construction of periodic solutions of nonlinear wave equations in higher dimension", *Geom. Funct. Anal.* **5**:4 (1995), 629–639. MR 96h:35011 Zbl 0834.35083

[Bourgain 1995b] J. Bourgain, "On the Cauchy problem for periodic KdV-type equations", pp. 17–86 in *Proceedings of the Conference in Honor of Jean-Pierre Kahane* (Orsay, 1993), CRC Press, Boca Raton, FL, 1995. Also available as a 1995 special issue of J. Fourier Anal. Appl. MR 96m:35270 Zbl 0891.35137

[Christ and Weinstein 1991] F. M. Christ and M. I. Weinstein, "Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation", *J. Funct. Anal.* **100**:1 (1991), 87–109. MR 92h:35203 Zbl 0743.35067

[Craig and Wayne 1993] W. Craig and C. E. Wayne, "Newton's method and periodic solutions of nonlinear wave equations", *Comm. Pure Appl. Math.* **46**:11 (1993), 1409–1498. MR 94m:35023 Zbl 0794.35104

[Dobrokhotov and Krichever 1991] S. Y. Dobrokhotov and I. M. Krichever, "Multiphase solutions of the Benjamin–Ono equation and their averaging", *Mat. Zametki* **49**:6 (1991), 42–58, 158. In Russian; translated in *Math. Notes* **49**:6, (1991), 583–594. MR 92g:35182 Zbl 0752.35058

[Dubrovin 1975] B. A. Dubrovin, "A periodic problem for the Korteweg-de Vries equation in a class of short-range potentials", *Funkcional. Anal. i Priložen.* **9**:3 (1975), 41–51. In Russian; translated in *Funct. Anal. Appl.* **9**:1 (1975), 41–42. MR 58 #6480 Zbl 0338.35022

[Düll 2012] W.-P. Düll, "Validity of the Korteweg-de Vries approximation for the two-dimensional water wave problem in the arc length formulation", *Comm. Pure Appl. Math.* **65**:3 (2012), 381–429. MR 2012m:35280 Zbl 1234.35218

[Erdoğan and Tzirakis 2013] M. B. Erdoğan and N. Tzirakis, "Global smoothing for the periodic KdV evolution", *Int. Math. Res. Not.* **2013**:20 (2013), 4589–4614. MR 3118870 Zbl 1295.35364

[Fibich et al. 2002] G. Fibich, B. Ilan, and G. Papanicolaou, "Self-focusing with fourth-order dispersion", *SIAM J. Appl. Math.* **62**:4 (2002), 1437–1462. MR 2003b:35198 Zbl 1003.35112

[Fibich et al. 2003] G. Fibich, B. Ilan, and S. Schochet, "Critical exponents and collapse of nonlinear Schrödinger equations with anisotropic fourth-order dispersion", *Nonlinearity* **16**:5 (2003), 1809–1821. MR 2004e:35209 Zbl 1040.35112

[Ghys 2007] É. Ghys, "Resonances and small divisors", pp. 187–213 in *Kolmogorov's heritage in mathematics*, edited by É. Charpentier et al., Springer, Berlin, 2007. MR 2009a:37116 Zbl 05294117

[Ginibre and Velo 1986] J. Ginibre and G. Velo, "Scattering theory in the energy space for a class of nonlinear Schrödinger equations", pp. 110–120 in *Semigroups, theory and applications, Vol. I* (Trieste, 1984), edited by H. Brezis et al., Pitman Res. Notes Math. Ser. **141**, Longman Sci. Tech., Harlow, 1986. MR 88a:35181 Zbl 0626.35073

[Hu and Li 2015] Y. Hu and X. Li, "Local well-posedness of periodic fifth-order KdV-type equations", *J. Geom. Anal.* **25**:2 (2015), 709–739. MR 3319948 Zbl 1321.35200

[Iooss et al. 2005] G. Iooss, P. I. Plotnikov, and J. F. Toland, "Standing waves on an infinitely deep perfect fluid under gravity", *Arch. Ration. Mech. Anal.* **177**:3 (2005), 367–478. MR 2007a:76017 Zbl 1176.76017

[Kappeler and Pöschel 2003] T. Kappeler and J. Pöschel, *KdV & KAM*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **45**, Springer, Berlin, 2003. MR 2004g:37099 Zbl 1032.37001

[Kappeler and Topalov 2006] T. Kappeler and P. Topalov, "Global wellposedness of KdV in $H^{-1}(\mathbb{T}, \mathbb{R})$", *Duke Math. J.* **135**:2 (2006), 327–360. MR 2007i:35199 Zbl 1106.35081

[Karpman and Shagalov 2000] V. I. Karpman and A. G. Shagalov, "Stability of solitons described by nonlinear Schrödinger-type equations with higher-order dispersion", *Phys. D* **144**:1-2 (2000), 194–210. MR 2001f:35333 Zbl 0962.35165

[Kato 1983] T. Kato, "On the Cauchy problem for the (generalized) Korteweg-de Vries equation", pp. 93–128 in *Studies in applied mathematics*, edited by V. Guillemin, Adv. Math. Suppl. Stud. **8**, Academic Press, New York, 1983. MR 86f:35160 Zbl 0549.34001

[Kenig et al. 1991] C. E. Kenig, G. Ponce, and L. Vega, "Oscillatory integrals and regularity of dispersive equations", *Indiana Univ. Math. J.* **40**:1 (1991), 33–69. MR 92d:35081 Zbl 0738.35022

[Kuksin 1988] S. B. Kuksin, "Perturbation theory of conditionally periodic solutions of infinite-dimensional Hamiltonian systems and its applications to the Korteweg-de Vries equation", *Mat. Sb.* (*N.S.*) **136(178)**:3 (1988), 396–412, 431. In Russian; translated in *Math. USSR-Sb.* **64**:2 (1989), 397–413. MR 89m:58096 Zbl 0657.58033

[Kuksin 1998] S. B. Kuksin, "A KAM-theorem for equations of the Korteweg-de Vries type", *Rev. Math. Math. Phys.* **10**:3 (1998), ii+64. MR 2001g:37140 Zbl 0920.35135

[Linares and Scialom 1993] F. Linares and M. Scialom, "On the smoothing properties of solutions to the modified Korteweg-de Vries equation", *J. Differential Equations* **106**:1 (1993), 141–154. MR 95b:35200 Zbl 0794.35117

[Liu 1997] Y. Liu, "Decay and scattering of small solutions of a generalized Boussinesq equation", *J. Funct. Anal.* **147**:1 (1997), 51–68. MR 98e:35144 Zbl 0884.35129

[de la Llave 2000] R. de la Llave, "Variational methods for quasi-periodic solutions of partial differential equations", pp. 214–228 in *Hamiltonian systems and celestial mechanics* (Pátzcuaro, 1998), edited by J. Delgado et al., World Sci. Monogr. Ser. Math. **6**, World Sci. Publ., 2000. MR 2001m:35012 Zbl 0987.35015

[Majda and Bertozzi 2002] A. J. Majda and A. L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Applied Mathematics **27**, Cambridge University Press, 2002. MR 2003a:76002 Zbl 0983.76001

[Matsuno 2004] Y. Matsuno, "New representations of multiperiodic and multisoliton solutions for a class of nonlocal soliton equations", *J. Phys. Soc. Japan* **73**:12 (2004), 3285–3293. MR 2005i:35237 Zbl 1066.35103

[Plotnikov and Toland 2001] P. I. Plotnikov and J. F. Toland, "Nash–Moser theory for standing water waves", *Arch. Ration. Mech. Anal.* **159**:1 (2001), 1–83. MR 2002k:76019 Zbl 1033.76005

[Ponce and Vega 1990] G. Ponce and L. Vega, "Nonlinear small data scattering for the generalized Korteweg-de Vries equation", *J. Funct. Anal.* **90**:2 (1990), 445–457. MR 91d:35191 Zbl 0771.35062

[Satsuma and Ishimori 1979] J. Satsuma and Y. Ishimori, "Periodic wave and rational soliton solutions of the Benjamin–Ono equation", *J. Phys. Soc. Japan* **46**:2 (1979), 681–687.

[Schneider and Wayne 2002] G. Schneider and C. E. Wayne, "The rigorous approximation of long-wavelength capillary-gravity waves", *Arch. Ration. Mech. Anal.* **162**:3 (2002), 247–285. MR 2003c:76020 Zbl 1055.76006

[Shatah 1985] J. Shatah, "Normal forms and quadratic nonlinear Klein–Gordon equations", *Comm. Pure Appl. Math.* **38**:5 (1985), 685–696. MR 87b:35160 Zbl 0597.35101

[Strauss 1974] W. A. Strauss, "Dispersion of low-energy waves for two conservative equations", *Arch. Rational Mech. Anal.* **55** (1974), 86–92. MR 50 #5230 Zbl 0289.35048

[Wayne 1990] C. E. Wayne, "Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory", *Comm. Math. Phys.* **127**:3 (1990), 479–528. MR 91b:58236 Zbl 0708.35087

[Wilkening 2008] J. Wilkening, "An infinite branching hierarchy of time-periodic solutions of the Benjamin–Ono equation", preprint, 2008. arXiv 0811.4209

Received 3 Oct 2014. Revised 31 Jul 2015. Accepted 13 Nov 2015.

DAVID M. AMBROSE: dma68@drexel.edu
Department of Mathematics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, United States

J. DOUGLAS WRIGHT: jdoug@math.drexel.edu
Department of Mathematics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, United States

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.43

THE BORDERLINES OF INVISIBILITY AND VISIBILITY IN CALDERÓN'S INVERSE PROBLEM

KARI ASTALA, MATTI LASSAS AND LASSI PÄIVÄRINTA

We consider the determination of a conductivity function in a two-dimensional domain from the Cauchy data of the solutions of the conductivity equation on the boundary. We prove uniqueness results for this inverse problem, posed by Calderón, for conductivities that are degenerate, that is, they may not be bounded from above or below. Elliptic equations with such coefficient functions are essential for physical models used in transformation optics and the study of metamaterials, e.g., the zero permittivity materials. In particular, we show that the elliptic inverse problems can be solved in a class of conductivities which is larger than L^∞. Also, we give new counterexamples for the uniqueness of the inverse conductivity problem.

We say that a conductivity is visible if the inverse problem is solvable so that the conductivity inside of the domain can be uniquely determined, up to a change of coordinates, using the boundary measurements. The original counterexamples for the inverse problem are related to the invisibility cloaking. This means that there are conductivities for which a part of the domain is shielded from detection via boundary measurements and even the existence of the shielded domain is hidden. Such conductivities are called invisibility cloaks.

In the present paper, we identify the borderline of the visible conductivities and the borderline of invisibility cloaking conductivities. Surprisingly, these borderlines are not the same. We show that between the visible and the cloaking conductivities, there are the electric holograms. These are conductivities which create an illusion of a nonexisting body. Such conductivities give counterexamples for the uniqueness of the inverse problem which are less degenerate than the previously known ones. These examples are constructed using transformation optics and the inverse maps of the optimal blow-up maps. The proofs of the uniqueness results for inverse problems are based on the complex geometrical optics and the Orlicz space techniques.

1. Introduction and main results 44
2. Proofs for the existence and uniqueness of the solution of the direct problem and for the counterexamples 56
3. Complex geometric optics solutions 67
4. Inverse conductivity problem with degenerate isotropic conductivity 73
5. Reduction of the inverse problem for an anisotropic conductivity to the isotropic case 91
Appendix: Orlicz spaces 94
Acknowledgements 95
References 95

MSC2010: 35R30.
Keywords: inverse conductivity problem, invisibility, quasiconformal mappings.

1. Introduction and main results

Invisibility cloaking has been a very topical subject in recent studies in mathematics, physics, and material science [Ammari et al. 2013; Alu and Engheta 2005; Greenleaf et al. 2007c; 2003a; 2003c; Milton and Nicorovici 2006; Leonhardt 2006; Liu 2013; Liu and Sun 2013; Pendry et al. 2006]. By invisibility cloaking we mean the possibility, both theoretical and practical, of shielding a region or object from detection via electromagnetic fields.

Counterexamples for inverse problems and the proposals for invisibility cloaking are closely related. In 2003, before the appearance of practical possibilities for cloaking, it was shown in [Greenleaf et al. 2003c] that passive objects can be coated with a layer of material with a degenerate conductivity which makes the object undetectable by electrostatic boundary measurements in such a way that the coated object appears in all measurements the same as a body made of homogeneous material. These constructions were based on blow-up maps and gave counterexamples for uniqueness of the inverse conductivity problem in the three- and higher-dimensional cases. In the two-dimensional case, the mathematical theory of the cloaking examples for the conductivity equation have been studied in [Kohn et al. 2008; 2010]. Besides for the conductivity equation, these results can be applied for other physical models based on elliptic equations.

The interest in cloaking was raised in particular in 2006 when it was realized that practical cloaking constructions are possible using so-called metamaterials which allow fairly arbitrary specification of electromagnetic material parameters. The construction of Leonhardt [2006] was based on conformal mapping on a nontrivial Riemannian surface. At the same time, Pendry et al. [2006] proposed a cloaking construction for Maxwell's equations using a blow-up map and the idea was demonstrated in laboratory experiments [Schurig et al. 2006]. Cloaking for the conductivity equation has been demonstrated in laboratory experiments by Yang et al. [2012]. In the now very large literature, there are also other suggestions for cloaking based on negative material parameters [Alu and Engheta 2005; Milton and Nicorovici 2006].

In this paper, we consider both new counterexamples and uniqueness results for inverse problems. We study Calderón's inverse problem in the two-dimensional case, that is, the question of whether an unknown conductivity distribution inside a domain can be determined from the voltage and current measurements made on the boundary of a simply connected domain $\Omega \subset \mathbb{R}^2$; see [Borcea 2002]. Mathematically the measurements correspond to the knowledge of the *Dirichlet-to-Neumann map* Λ_σ (or the quadratic form) associated to σ, i.e., the map taking the Dirichlet boundary values of the solution of the conductivity equation

$$\nabla \cdot \sigma(x) \nabla u(x) = 0 \quad \text{in } \Omega \tag{1-1}$$

to the corresponding Neumann boundary values,

$$\Lambda_\sigma : u|_{\partial\Omega} \mapsto \nu \cdot \sigma \nabla u|_{\partial\Omega}. \tag{1-2}$$

In the classical theory of the problem, the conductivity σ is bounded uniformly from above and below. The problem was originally proposed by Calderón [1980]. Sylvester and Uhlmann [1987] proved the unique identifiability of the conductivity in dimensions three and higher for isotropic conductivities which are C^∞-smooth, and Nachman [1988] gave a reconstruction method. In three dimensions or higher, unique identifiability of the conductivity is proven in [Haberman and Tataru 2013] for C^1-conductivities;

for earlier studies on the topic, see [Greenleaf et al. 2003b; Päivärinta et al. 2003]. The problem has also been solved with measurements only on a part of the boundary [Kenig et al. 2007].

In two dimensions, the first global solution of the inverse conductivity problem is due to Nachman [1996] for conductivities with two derivatives. In this seminal paper, the $\bar{\partial}$ technique was used for the first time in the study of Calderón's inverse problem. The smoothness requirements were reduced in [Brown and Uhlmann 1997] to Lipschitz conductivities. Finally, in [Astala and Päivärinta 2006] the uniqueness of the inverse problem was proven in the form that the problem was originally formulated in [Calderón 1980], i.e., for general isotropic conductivities in L^∞ which are bounded from below and above by positive constants. Stability of this reconstruction is studied in [Alessandrini 1988; Barceló et al. 2001; 2007] and the numerical solutions in [Astala et al. 2011; Isaacson et al. 2004; Knudsen et al. 2007; 2009; Mueller and Siltanen 2012; Siltanen et al. 2000].

The Calderón problem with anisotropic, i.e., matrix-valued, conductivities that are uniformly bounded from above and below, has been studied in two dimensions [Sylvester 1990; Nachman 1996; Lassas et al. 2003; Lassas and Uhlmann 2001; Astala et al. 2005; Imanuvilov et al. 2010] and in dimensions $n \geq 3$ [Lee and Uhlmann 1989; Lassas and Uhlmann 2001; Dos Santos Ferreira et al. 2009]. For example, for the anisotropic inverse conductivity problem in the two-dimensional case, it is known that the Dirichlet-to-Neumann map determines a regular conductivity tensor only up to a diffeomorphism $F : \bar{\Omega} \to \bar{\Omega}$; i.e., one can obtain an image of the interior of Ω in deformed coordinates. This implies that the inverse problem is not uniquely solvable, but the nonuniqueness of the problem can be characterized. This makes it possible, e.g., to find the unique conductivity that is closest to isotropic ones [Kolehmainen et al. 2005; 2010; 2013]. We note that this problem in higher dimensions is presently solved only in special cases, when the conductivity is real analytic; see [Lassas et al. 2003; Lassas and Uhlmann 2001].

In this work, we will study the inverse conductivity problem in the two-dimensional case with degenerate conductivities. Such conductivities appear in physical models where the medium varies continuously from a perfect conductor to a perfect insulator or in high-contrast problems [Borcea et al. 1996; Borcea 1999]. As an example, we may consider a case where the conductivity goes to zero or to infinity near ∂D, where $D \subset \Omega$ is a smooth open set. We ask what kind of degeneracy prevents solving the inverse problem; that is, we study what is the border of visibility. We also ask what kind of degeneracy makes it possible to coat an arbitrary object so that it appears the same as a homogeneous body in all static measurements; that is, we study what is the border of the invisibility cloaking. *Surprisingly, these borders are not the same.* We identify these borderlines and show that between them there are the *electric holograms*, that is, the conductivities creating an illusion of a *nonexistent* body (see Figure 1). These conductivities are counterexamples for the unique solvability of inverse problems for which even the topology of the domain cannot be determined using boundary measurements. Our main results for the uniqueness of the inverse problem are given in Theorems 1.8, 1.9, and 1.11 and the counterexamples are formulated in Theorems 1.6 and 1.7.

The cloaking constructions have given rise to the design technique called *transformation optics*. The metamaterials built to operate at microwave frequencies [Schurig et al. 2006] and near the optical frequencies [Ergin et al. 2010] are inherently prone to dispersion, so that realistic cloaking must currently be considered as occurring at a very narrow range of wavelengths. Fortunately, in many physical applications

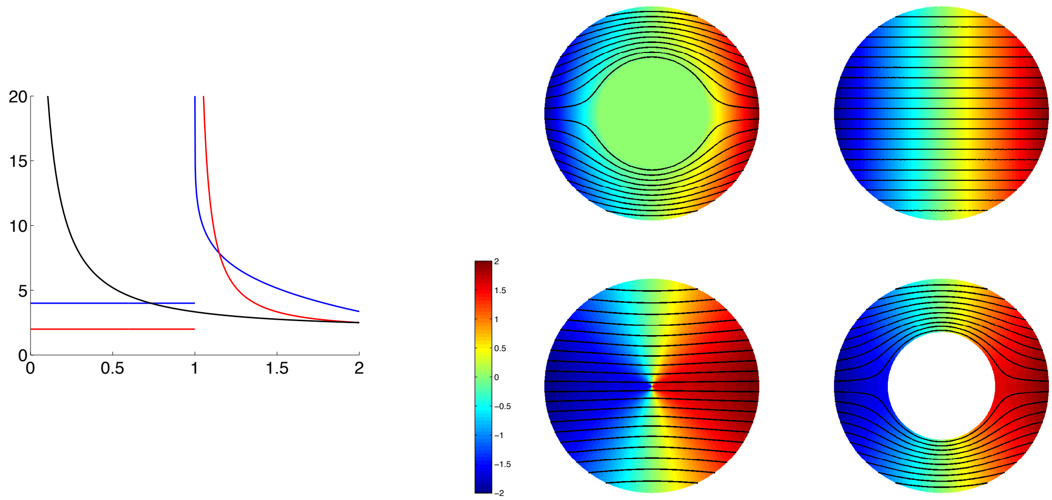

Figure 1. **Left:** $\text{tr}(\sigma)$ of three radial and singular conductivities on the positive x axis. The curves correspond to the invisibility cloaking conductivity (red), with the singularity $\sigma^{22}(x,0) \sim (|x|-1)^{-1}$ for $|x|>1$, a visible conductivity (blue) with a $\log\log$-type singularity at $|x|=1$, and an electric hologram (black) with the conductivity having the singularity $\sigma^{11}(x,0) \sim |x|^{-1}$. **Right, top:** All measurements on the boundary of the invisibility cloak (left) coincide with the measurements for the homogeneous disc (right). The color shows the value of the solution u with the boundary value $u(x,y)|_{\partial B(2)} = x$ and the black curves are the integral curves of the current $-\sigma\nabla u$. **Right, bottom:** All measurements on the boundary of the electric hologram (left) coincide with the measurements for an isolating disc covered with the homogeneous medium (right). The solutions and the current lines corresponding to the boundary value $u|_{\partial B(2)} = x$ are shown.

the materials need to operate only near a single frequency. The cloaking-type constructions have also inspired suggestions for possible devices producing extreme effects on wave propagation, including invisibility cloaks for magnetostatics [Gömöry et al. 2012], acoustics [Chen and Chan 2007a], quantum mechanics [Greenleaf et al. 2007a, 2008; 2011a], field rotators [Chen and Chan 2007b], electromagnetic wormholes [Greenleaf et al. 2007b], invisible sensors [Alu and Engheta 2009; Greenleaf et al. 2011b], perfect absorbers [Landy et al. 2008], and cloaked wave amplifiers [Greenleaf et al. 2012]. We also note that the differential equations with degenerate coefficients modeling cloaking devices have turned out to have interesting properties, such as nonexistence results for solutions with nonzero sources and local [Greenleaf et al. 2007c; Liu and Zhou 2011] and nonlocal [Lassas and Zhou 2011; Nguyen 2012] hidden boundary conditions. For reviews on the topic, see [Greenleaf et al. 2009a; 2009b].

Finally, the classical assumption that the electromagnetic material parameters (i.e., the coefficient functions in the elliptic equations) are uniformly bounded from below by positive constants is not valid in the modern study of materials, e.g., on the optical frequencies [Capolino 2009]. Thus one of the aims of this paper is to bring the mathematical study of elliptic equations and inverse problems closer to current topics in optics and imaging methods in physics.

The structure of the paper is the following. The main results and the formulation of the boundary measurements are presented in the first section. The proofs for the existence of the solutions of the direct problem as well as for the new counterexamples and the invisibility cloaking examples with a nonsmooth background are given in Section 2. The uniqueness result for isotropic conductivities is proven in Sections 3–4 and the reduction of the general problem to the isotropic case is shown in Section 5. In Sections 3–5, the degeneracy of the conductivity yields that the exponentially growing solutions, the standard tools used to study Calderón's inverse problem, cannot be constructed using purely microlocal or functional analytic methods. Because of this, we will extensively need the topological properties of the solutions: By Stoilow's theorem, the solutions are compositions of analytic functions and homeomorphisms. Using this, the continuity properties of the weakly monotone maps, and Orlicz estimates holding for homeomorphisms, we prove the existence of the solutions in the Sobolev–Orlicz spaces. These spaces are chosen so that we can obtain the subexponential asymptotics for the families of exponentially growing solutions needed in the $\bar{\partial}$ technique used to solve the inverse problem.

1A. ***Definition of measurements and solvability.*** Let $\Omega \subset \mathbb{R}^2$ be a bounded simply connected domain with a C^∞-smooth boundary. Let $\Sigma = \Sigma(\Omega)$ be the class of measurable matrix-valued functions $\sigma : \Omega \to M$, where M is the set of generalized matrices $\boldsymbol{m}$ of the form

$$\boldsymbol{m} = U \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U^t,$$

where $U \in \mathbb{R}^{2\times 2}$ is an orthogonal matrix, $U^{-1} = U^t$ and $\lambda_1, \lambda_2 \in [0, \infty)$. We denote by $W^{s,p}(\Omega)$ and $H^s(\Omega) = W^{s,2}(\Omega)$ the standard Sobolev spaces.

In the following, let $dm(z)$ denote the Lebesgue measure in $\mathbb{C}$ and $|E|$ be the Lebesgue measure of the set $E \subset \mathbb{C}$. Instead of defining the Dirichlet-to-Neumann operator for the above conductivities, we consider the corresponding quadratic forms.

Definition 1.1. Let $h \in H^{1/2}(\partial\Omega)$. The *Dirichlet-to-Neumann quadratic form* corresponding to the conductivity $\sigma \in \Sigma(\Omega)$ is given by

$$L_\sigma[h] = \inf A_\sigma[u], \quad \text{where } A_\sigma[u] = \int_\Omega \sigma(z)\nabla u(z) \cdot \nabla u(z)\, dm(z), \tag{1-3}$$

and the infimum is taken over real-valued $u \in L^1(\Omega)$ such that $\nabla u \in L^1(\Omega; \mathbb{R}^2)$ and $u|_{\partial\Omega} = h$. In the case where $L_\sigma[h] < \infty$ and $A_\sigma[u]$ reaches its minimum at some u, we say that u is a $W^{1,1}(\Omega)$ solution of the conductivity problem.

When σ is smooth and bounded from below and above by positive constants, $L_\sigma[h]$ is the quadratic form corresponding the Dirichlet-to-Neumann map (1-2),

$$L_\sigma[h] = \int_{\partial\Omega} h \Lambda_\sigma h\, dS, \tag{1-4}$$

where dS is the length measure on $\partial\Omega$. Physically, $L_\sigma[h]$ corresponds to the power needed to keep voltage h at the boundary. For smooth conductivities bounded from below, for every $h \in H^{1/2}(\partial\Omega)$, the integral $A_\sigma[u]$ always has a unique minimizer $u \in H^1(\Omega)$ with $u|_{\partial\Omega} = h$, which is also a distributional solution to (1-1). Conversely, for functions $u \in H^1(\Omega)$, their traces lie in $H^{1/2}(\partial\Omega)$. It is for this reason that we chose to consider the $H^{1/2}$-boundary functions also in the most general case. We interpret that the Dirichlet-to-Neumann form corresponds to an idealization of the boundary measurements for $\sigma \in \Sigma(\Omega)$.

We note that the conductivities studied in the context of cloaking are not even in L^1_{loc}. As σ is unbounded, it is possible that $L_\sigma[h] = \infty$. Even if $L_\sigma[h]$ is finite, the minimization problem in (1-3) may generally have no minimizer and even if they exist, the minimizers need not be distributional solutions to (1-1). However, if the singularities of σ are not too strong, minimizers satisfying (1-1) do always exist. To show this, we need to define suitable subclasses of degenerate conductivities.

Let $\sigma \in \Sigma(\Omega)$. We start with precise quantities describing the possible degeneracy or loss of uniform ellipticity. First, a natural measure of the anisotropy of the conductivity σ at $z \in \Omega$ is

$$k_\sigma(z) = \sqrt{\frac{\lambda_1(z)}{\lambda_2(z)}},$$

where $\lambda_1(z)$ and $\lambda_2(z)$ are the eigenvalues of the matrix $\sigma(z)$ with $\lambda_1(z) \geq \lambda_2(z)$. If we want to simultaneously control both the size and the anisotropy, this is measured by the *ellipticity* $K(z) := K_\sigma(z)$ of $\sigma(z)$, i.e., the smallest number $1 \leq K(z) \leq \infty$ such that

$$\frac{1}{K(z)}|\xi|^2 \leq \xi \cdot \sigma(z)\xi \leq K(z)|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^2. \tag{1-5}$$

For a general, positive matrix-valued function $\sigma(z)$, we have at $z \in \Omega$ that

$$K(z) = k_\sigma(z) \max\big\{(\det \sigma(z))^{1/2}, (\det \sigma(z))^{-1/2}\big\}. \tag{1-6}$$

Consequently, we always have the following simple estimates.

Lemma 1.2. *For any measurable matrix function $\sigma(z)$, we have*

$$\tfrac{1}{4}\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1})\big) \leq K(z) \leq \operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1}).$$

Proof. Let $\lambda_{\max}$ and $\lambda_{\min}$ be the eigenvalues of $\sigma = \sigma(z)$. Then $K(z) = \max(\lambda_{\max}, \lambda_{\min}^{-1})$. Since $\operatorname{tr}\sigma(z) = \lambda_{\max} + \lambda_{\min}$ and $\operatorname{tr}(\sigma(z)^{-1}) = \lambda_{\max}^{-1} + \lambda_{\min}^{-1}$, the claim follows. □

Due to Lemma 1.2, we use the quantity $\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1})$ as a measure of size and anisotropy of $\sigma(z)$.

For the degenerate elliptic equations, it may be that the optimization problem (1-3) has a minimizer which satisfies the conductivity equation but this solution may not have the standard $W^{1,2}_{\mathrm{loc}}$ regularity. Therefore more subtle smoothness estimates are required. We start with the exponentially integrable conductivities, and the natural energy estimates they require. As an important consequence, we will see the correct Sobolev–Orlicz regularity to work with. These observations are based on the following elementary inequality.

Lemma 1.3. *Let $K \geq 1$ and $A \in \mathbb{R}^{2\times 2}$ be a symmetric matrix satisfying*

$$\frac{1}{K}|\xi|^2 \leq \xi \cdot A\xi \leq K|\xi|^2, \quad \xi \in \mathbb{R}^2.$$

Then for every $p > 0$,

$$\frac{|\xi|^2}{\log(e+|\xi|^2)} + \frac{|A\xi|^2}{\log(e+|A\xi|^2)} \le \frac{2}{p}(\xi \cdot A\xi + e^{pK}).$$

Proof. Since $K \ge 1$ and $t \mapsto t/\log(e+t)$ is an increasing function, we have

$$\begin{aligned}\frac{|\xi|^2}{\log(e+|\xi|^2)} &\le \frac{K\xi \cdot A\xi}{\log(e+K\xi \cdot A\xi)} \\ &\le \frac{1}{p}\left(\frac{\xi \cdot A\xi}{\log(e+\xi \cdot A\xi)}\right)pK \\ &\le \frac{1}{p}(\xi \cdot A\xi + e^{pK}),\end{aligned}$$

where the last estimate follows from the inequality

$$ab \le a\log(e+a) + e^b, \quad a, b \ge 0.$$

Moreover, as K is at least as large as the maximal eigenvalue of A, we have $|A\xi|^2 \le K\xi \cdot A\xi$. Thus we see as above that

$$\frac{|A\xi|^2}{\log(e+|A\xi|^2)} \le \frac{K\xi \cdot A\xi}{\log(e+K\xi \cdot A\xi)} \le \frac{1}{p}(\xi \cdot A\xi + e^{pK}).$$

Adding the above estimates together proves the claim. □

Lemma 1.3 implies in particular that if $\sigma(z)$ is symmetric-matrix-valued function satisfying (1-5) for a.e. $z \in \Omega$ and $u \in W^{1,1}(\Omega)$, then we always have

$$\begin{aligned} p\int_\Omega \frac{|\nabla u(z)|^2}{\log(e+|\nabla u(z)|^2)}\,dm(z) &\le \int_\Omega \nabla u(z)\cdot\sigma(z)\nabla u(z)\,dm(z) + \int_\Omega e^{pK(z)}\,dm(z), \\ p\int_\Omega \frac{|\sigma(z)\nabla u(z)|^2}{\log(e+|\sigma(z)\nabla u(z)|^2)}\,dm(z) &\le \int_\Omega \nabla u(z)\cdot\sigma(z)\nabla u(z)\,dm(z) + \int_\Omega e^{pK(z)}\,dm(z). \end{aligned} \tag{1-7}$$

Note that these inequalities are valid whether u is a solution of the conductivity equation or not!

Due to (1-7), we see that to analyze finite energy solutions corresponding to a singular conductivity of exponentially integrable ellipticity, we are naturally led to consider the regularity gauge

$$Q(t) = \frac{t^2}{\log(e+t)}, \quad t \ge 0. \tag{1-8}$$

We say accordingly that f belongs to the Orlicz space $W^{1,Q}(\Omega)$ (see the Appendix) if f and its first distributional derivatives are in $L^1(\Omega)$ and

$$\int_\Omega \frac{|\nabla f(z)|^2}{\log(e+|\nabla f(z)|)}\,dm(z) < \infty.$$

The first existence result for solutions corresponding to degenerate conductivities is given as follows.

Theorem 1.4. *Let $\sigma(z)$ be a measurable symmetric-matrix-valued function. Suppose further that for some $p > 0$,*

$$\int_\Omega \exp\big(p\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1})\big)\big)\, dm(z) = C_1 < \infty. \tag{1-9}$$

Then, if $h \in H^{1/2}(\partial\Omega)$ is such that $L_\sigma[h] < \infty$ and there is a unique $w \in W^{1,1}(\Omega)$, $w|_{\partial\Omega} = h$ such that

$$A_\sigma[w] = \inf\big\{A_\sigma[v] : v \in W^{1,1}(\Omega),\, v|_{\partial\Omega} = h\big\}. \tag{1-10}$$

Moreover, w satisfies the conductivity equation

$$\nabla\cdot\sigma\nabla w = 0 \quad \text{in } \Omega \tag{1-11}$$

in the sense of distributions, and it has the regularity $w \in W^{1,Q}(\Omega) \cap C(\Omega)$.

Note that if σ is bounded near $\partial\Omega$ then $L_\sigma[h] < \infty$ for all $h \in H^{1/2}(\partial\Omega)$. Theorem 1.4 is proven in Theorem 2.1 and Corollary 2.3 in a more general setting.

Theorem 1.4 yields that for conductivities satisfying (1-9) and equal to 1 near $\partial\Omega$, we can define the Dirichlet-to-Neumann map

$$\Lambda_\sigma : H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega), \quad \Lambda_\sigma(u|_{\partial\Omega}) = \nu\cdot\sigma\nabla u|_{\partial\Omega}, \tag{1-12}$$

where u satisfies (1-1). Many inverse scattering problems (see [Colton and Kress 2013]) can also be formulated in terms of Λ_σ.

The reader should consider the exponential condition (1-9) as being close to the optimal one, still allowing uniqueness in the inverse problem. Indeed, in view of Theorem 1.7 and Section 1E below, the most general situation where the Calderón inverse problem can be solved involves conductivities whose singularities satisfy a physically interesting small relaxation of the condition (1-9). Before solving inverse problems for conductivities satisfying (1-9), we discuss some counterexamples.

1B. ***Counterexamples for the unique solvability of the inverse problem.*** Let $F : \Omega_1 \to \Omega_2$, $y = F(x)$, be an orientation-preserving homeomorphism between domains $\Omega_1, \Omega_2 \subset \mathbb{C}$ for which F and its inverse F^{-1} are at least $W^{1,1}$-smooth and let $\sigma(x) = [\sigma^{jk}(x)]_{j,k=1}^2 \in \Sigma(\Omega_1)$ be a conductivity on Ω_1. Then the map F pushes σ forward to a conductivity $(F_*\sigma)(y)$, defined on Ω_2 and given by

$$(F_*\sigma)(y) = \frac{1}{[\det DF(x)]} DF(x)\sigma(x)\, DF(x)^t, \quad x = F^{-1}(y). \tag{1-13}$$

The main methods for constructing counterexamples to Calderón's problem are based on the following principle.

Proposition 1.5. *Assume that $\sigma, \tilde\sigma \in \Sigma(\Omega)$ satisfy (1-9), and let $F : \Omega \to \widetilde\Omega$ be a homeomorphism so that F and F^{-1} are $W^{1,Q}$-smooth and C^1-smooth near the boundary, and $F|_{\partial\Omega} = \mathrm{id}$. Suppose that $\tilde\sigma = F_*\sigma$. Then $L_\sigma = L_{\tilde\sigma}$.*

This proposition generalizes the observation of L. Tartar expanded upon in [Kohn and Vogelius 1984] to less smooth diffeomorphisms and conductivities and it follows from Lemma 2.4 proven later.

1C. ***Counterexample 1: invisibility cloaking.*** We consider here invisibility cloaking in a general background σ; that is, we aim to coat an arbitrary body with a layer of exotic material so that the coated body appears in measurements the same as the background conductivity σ. Usually one is interested in the case when the background conductivity σ is constant, isotropic, and $\sigma = 1 \cdot I$. However, we consider here a more general case and assume that σ is an L^∞-smooth conductivity in $\overline{B}(R) = \overline{B(R)}$, with $R = 2$, $\sigma(z) \geq c_0 I$, $c_0 > 0$. Here, $B(\rho)$ is an open two-dimensional disc of radius ρ and center zero and $\overline{B}(\rho)$ is its closure. Consider a homeomorphism

$$F : \overline{B}(2) \setminus \{0\} \to \overline{B}(2) \setminus \mathcal{K}, \tag{1-14}$$

where $\mathcal{K} \subset B(2)$ is a compact set which is the closure of a smooth open set and suppose F and its inverse F^{-1} are C^1-smooth in $\overline{B}(2) \setminus \{0\}$ and $\overline{B}(2) \setminus \mathcal{K}$, correspondingly. We also require that $F(z) = z$ for $z \in \partial B(2)$. The standard example of invisibility cloaking [Greenleaf et al. 2003c; Pendry et al. 2006] is the case when $\mathcal{K} = \overline{B}(1)$ and the map is given by

$$F_0(z) = \left(\frac{|z|}{2} + 1\right)\frac{z}{|z|}. \tag{1-15}$$

Using the map (1-14), we define a singular conductivity

$$\tilde{\sigma}(z) = \begin{cases} (F_*\sigma)(z) & \text{for } z \in B(2) \setminus \mathcal{K}, \\ \eta(z) & \text{for } z \in \mathcal{K}, \end{cases} \tag{1-16}$$

where $\eta(z) = [\eta^{jk}(x)]$ is any symmetric measurable matrix on $\mathcal{K}$ satisfying $c_1 I \leq \eta(z) \leq c_2 I$ with $c_1, c_2 > 0$. The conductivity $\tilde{\sigma}$ is called the cloaking conductivity obtained from the transformation map F and background conductivity σ and $\eta(z)$ is the conductivity of the cloaked (i.e., hidden) object.

In particular, choosing σ to be the constant conductivity $\sigma = 1$, $\mathcal{K} = \overline{B}(1)$, and F to be the map F_0 given in (1-15), we obtain the standard example of the invisibility cloaking. In dimensions $n \geq 3$, it was shown in [Greenleaf et al. 2003c] that the Dirichlet-to-Neumann map corresponding to $H^1(\Omega)$-solutions for the conductivity (1-16) coincide with the Dirichlet-to-Neumann map for $\sigma = 1$. In 2008, the analogous result was proven in the two-dimensional case in [Kohn et al. 2008]. For cloaking results for the Helmholtz equation with frequency $k \neq 0$ and for Maxwell's system in dimensions $n \geq 3$, see results in [Greenleaf et al. 2007c]. We note also that John Ball [1982] has used the push-forward by the analogous radial blow-up maps to study the discontinuity of solutions of partial differential equations, in particular the appearance of cavitation in nonlinear elasticity.

In the sequel, we consider cloaking results using measurements given in Definition 1.1. As we have formulated the boundary measurements in a new way, that is, in terms of the Dirichlet-to-Neumann forms L_σ associated to the class $W^{1,1}(\Omega)$, we present in Section 2D the complete proof of the following proposition, extending [Greenleaf et al. 2003c, Theorem 3]:

Theorem 1.6. (i) *Let $\sigma \in L^\infty(B(2))$ be a scalar conductivity, $\sigma(x) \geq c_0 > 0$, $\mathcal{K} \subset B(2)$ be a relatively compact open set with smooth boundary and*

$$F : \overline{B}(2) \setminus \{0\} \to \overline{B}(2) \setminus \mathcal{K}$$

be a homeomorphism. Assume that F and F^{-1} are C^1-smooth in $\bar{B}(2)\setminus\{0\}$ and $\bar{B}(2)\setminus\mathcal{K}$, correspondingly, and $F|_{\partial B(2)} = \mathrm{id}$. Moreover, assume there is $C_0 > 0$ such that

$$\|DF^{-1}(x)\| \le C_0 \quad \text{for all } x \in \bar{B}(2)\setminus\mathcal{K}.$$

Let $\tilde{\sigma}$ be the conductivity defined in (1-16). *Then the boundary measurements for $\tilde{\sigma}$ and σ coincide in the sense that $L_{\tilde{\sigma}} = L_\sigma$.*

(ii) *Let $\tilde{\sigma}$ be a cloaking conductivity of the form* (1-16) *obtained from the transformation map F and the background conductivity σ, where F and σ satisfy the conditions in* (i). *Then*

$$\mathrm{tr}(\tilde{\sigma}) \notin L^1(B(2)\setminus\mathcal{K}). \tag{1-17}$$

The result (1-17) is optimal in the following sense. When F is the map F_0 in (1-15) and $\sigma = 1$, the eigenvalues of the cloaking conductivity $\tilde{\sigma}$ in $B(2)\setminus\bar{B}(1)$ behave asymptotically as $|z|-1$ and $(|z|-1)^{-1}$ as $|z| \to 1$. This cloaking conductivity has so strong a degeneracy that (1-17) holds. On the other hand,

$$\mathrm{tr}(\tilde{\sigma}) \in L^1_{\mathrm{weak}}(B(2)), \tag{1-18}$$

where L^1_{weak} is the weak-L^1 space. We note that in the case when $\sigma = 1$, $\det(\tilde{\sigma})$ is identically 1 in $B(2)\setminus\bar{B}(1)$.

The formula (1-18) for the blow-up map F_0 in (1-15) and Theorem 1.6 identify the *borderline of the invisibility* for the trace of the conductivity: Any cloaking conductivity $\tilde{\sigma}$ satisfies $\mathrm{tr}(\tilde{\sigma}) \notin L^1(B(2))$ and there is an example of a cloaking conductivity for which $\mathrm{tr}(\tilde{\sigma}) \in L^1_{\mathrm{weak}}(B(2))$. Thus the borderline of invisibility is the same as the border between the space L^1 and the weak-L^1 space.

1D. ***Counterexample 2: illusion of a nonexistent obstacle.*** Next we consider new counterexamples for the inverse problem which could be considered as creating an illusion of a nonexisting obstacle. The example is based on a radial shrinking map, that is, a mapping $B(2)\setminus\bar{B}(1) \to B(2)\setminus\{0\}$. The suitable maps are the inverse maps of the blow-up maps $F_1 : B(2)\setminus\{0\} \to B(2)\setminus\bar{B}(1)$, which are constructed by Iwaniec and Martin [2001] and have the optimal smoothness. Alternative constructions for such blow-up maps have also been proposed by Kauhanen et al. [2003]. Using the properties of these maps and defining a conductivity $\sigma_1 = (F_1^{-1})_* 1$ on $B(2)\setminus\{0\}$, we will later prove the following result.

Theorem 1.7. *Let γ_1 be a conductivity in $B(2)$ which is identically* 1 *in $B(2)\setminus\bar{B}(1)$ and zero in $B(1)$ and $\mathcal{A} : [1,\infty] \to [0,\infty]$ be any strictly increasing positive smooth function with $\mathcal{A}(1) = 0$ which is sublinear in the sense that*

$$\int_1^\infty \frac{\mathcal{A}(t)}{t^2}\,dt < \infty. \tag{1-19}$$

Then there is a conductivity $\sigma_1 \in \Sigma(B_2)$ satisfying $\det(\sigma_1) = 1$ and

$$\int_{B(2)} \exp\big(\mathcal{A}\big(\mathrm{tr}(\sigma_1) + \mathrm{tr}(\sigma_1^{-1})\big)\big)\,dm(z) < \infty \tag{1-20}$$

such that $L_{\sigma_1} = L_{\gamma_1}$, i.e., the boundary measurements corresponding to σ_1 and γ_1 coincide.

We observe that, for instance, the function $\mathcal{A}_0(t) = t/(1+\log t)^{1+\varepsilon}$ satisfies (1-19) and for such a weight function, we have $\sigma_1 \in L^1(B_2)$. The proof of Theorem 1.7 is given in Section 2D.

Note that γ_1 corresponds to the case when $B(1)$ is a perfect insulator which is surrounded with constant conductivity 1. Thus Theorem 1.7 can be interpreted by saying that there is a relatively weakly degenerated conductivity satisfying integrability condition (1-20) that creates in the boundary observations an illusion of an obstacle that does not exist (see [Lai et al. 2009] for related results based on use of negative medium). Thus the conductivity can be considered as "electric hologram". As the obstacle can be considered as a "hole" in the domain, we can say also that even the topology of the domain cannot be detected. In other words, Calderón's program to image the conductivity inside a domain using the boundary measurements cannot work within the class of degenerate conductivities satisfying (1-19) and (1-20).

1E. ***Positive results for Calderón's inverse problem.*** Let us formulate our first main theorem which deals with inverse problems for anisotropic conductivities where both the trace and the determinant of the conductivity can be degenerate.

Theorem 1.8. *Let $\Omega \subset \mathbb{C}$ be a bounded simply connected domain with smooth boundary. Let $\sigma_1, \sigma_2 \in \Sigma(\Omega)$ be matrix-valued conductivities in Ω which satisfy the integrability condition*

$$\int_\Omega \exp\big(p\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1})\big)\big)\, dm(z) < \infty$$

for some $p > 1$. Moreover, assume that

$$\int_\Omega \mathcal{E}(q \det \sigma_j(z))\, dm(z) < \infty \quad \textit{for some } q > 0, \tag{1-21}$$

where $\mathcal{E}(t) = \exp(\exp(\exp(t^{1/2}+t^{-1/2})))$ and $L_{\sigma_1} = L_{\sigma_2}$. Then there is a $W^{1,1}_{\mathrm{loc}}$-homeomorphism $F : \Omega \to \Omega$ satisfying $F|_{\partial\Omega} = \mathrm{id}$ such that

$$\sigma_1 = F_* \sigma_2. \tag{1-22}$$

Equation (1-22) can be stated as saying that σ_1 and σ_2 are the same up to a change of coordinates; that is, the underlying manifold structures corresponding to these conductivities are the same; see [Lee and Uhlmann 1989; Lassas and Uhlmann 2001].

In the case when the conductivities are isotropic, we can improve the result of Theorem 1.8. The following theorem is our second main result for uniqueness of the inverse problem. For the earlier conjectures on the problem, see [Ingerman 2000].

Theorem 1.9. *Let $\Omega \subset \mathbb{C}$ be a bounded simply connected domain with smooth boundary. If $\sigma_1, \sigma_2 \in \Sigma(\Omega)$ are isotropic conductivities, i.e., $\sigma_j(z) = \gamma_j(z) I$, $\gamma_j(z) \in [0, \infty]$ satisfying*

$$\int_\Omega \exp\Big(\exp\Big(q\Big(\gamma_j(z) + \frac{1}{\gamma_j(z)}\Big)\Big)\Big)\, dm(z) < \infty \quad \textit{for some } q > 0, \tag{1-23}$$

and $L_{\sigma_1} = L_{\sigma_2}$, then $\sigma_1 = \sigma_2$.

Let us next consider anisotropic conductivities with bounded determinant but more degenerate ellipticity function $K_\sigma(z)$ defined in (1-5), and ask how far can we then generalize Theorem 1.8. Motivated by the counterexample given in Theorem 1.7, we consider the following class: we say that $\sigma \in \Sigma(\Omega)$ has an (at most) *exponentially degenerated anisotropy* with a weight $\mathcal{A}$, denoted $\sigma \in \Sigma_{\mathcal{A}} := \Sigma_{\mathcal{A}}(\Omega)$, if $\sigma(z) \in \mathbb{R}^{2\times 2}$ for a.e. $z \in \Omega$ and

$$\int_\Omega \exp\bigl(\mathcal{A}\bigl(\operatorname{tr}\sigma + \operatorname{tr}(\sigma^{-1})\bigr)\bigr)\, dm(z) < \infty. \tag{1-24}$$

In view of Theorem 1.7, for obtaining uniqueness for the inverse problem, we need to consider weights that are strictly increasing positive smooth functions $\mathcal{A} : [1, \infty] \to [0, \infty]$, $\mathcal{A}(1) = 0$, with

$$\int_1^\infty \frac{\mathcal{A}(t)}{t^2}\, dt = \infty \qquad \text{and} \qquad t\mathcal{A}'(t) \to \infty \quad \text{as } t \to \infty. \tag{1-25}$$

We say that $\mathcal{A}$ has *almost linear growth* if (1-25) holds. The point here is the first condition, that is, the divergence of the integral. The second condition is a technicality, which is satisfied by all weights one encounters in practice (which do not oscillate too much); the condition guarantees that the Sobolev-gauge function $P(t)$ defined below in (1-26) is equivalent to a convex function for large t; see [Astala et al. 2009, Lemma 20.5.4].

Note, in particular, that affine weights $\mathcal{A}(t) = pt - p$, $p > 0$, satisfy the condition (1-25). To develop uniqueness results for inverse problems within the class $\Sigma_{\mathcal{A}}$, the first problems we face are to establish the right Sobolev–Orlicz regularity for the solutions u of finite energy, $A_\sigma[u] < \infty$, and to solve the Dirichlet problem with given boundary values.

To start with this, we need the counterpart of the gauge $Q(t)$ defined in (1-8). In the case of a general weight $\mathcal{A}$, we define

$$P(t) = \begin{cases} t^2 & \text{for } 0 \leqslant t < 1, \\ \dfrac{t^2}{\mathcal{A}^{-1}(\log(t^2))} & \text{for } t \geqslant 1, \end{cases} \tag{1-26}$$

where $\mathcal{A}^{-1}$ is the inverse function of $\mathcal{A}$. As an example, note that if $\mathcal{A}$ is affine, $\mathcal{A}(t) = pt - p$ for some number $p > 0$, then the condition (1-24) takes us back to the exponentially integrable distortion of Theorem 1.8, while $P(t) = t^2\bigl(1 + (1/p)\log^+(t^2)\bigr)^{-1}$ is equivalent to the gauge function $Q(t)$ used in (1-8).

The inequalities (1-7) corresponding to the case when $\mathcal{A}$ is affine can be generalized for the following result holding for general gauge $\mathcal{A}$ satisfying (1-25).

Lemma 1.10. *Suppose $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$ and $\mathcal{A}$ satisfies the almost linear growth condition* (1-25). *Then*

$$\int_\Omega \bigl(P(|\nabla u|) + P(|\sigma\nabla u|)\bigr)\, dm \le 2\int_\Omega e^{\mathcal{A}(\operatorname{tr}\sigma + \operatorname{tr}(\sigma^{-1}))}\, dm(z) + 2\int_\Omega \nabla u \cdot \sigma \nabla u\, dm$$

for every measurable function of symmetric matrices $\sigma(z) \in \mathbb{R}^{2\times 2}$.

Proof. We have, in fact, pointwise estimates. For these, note first that the conditions for $\mathcal{A}(t)$ imply that $P(t) \le t^2$ for every $t \ge 0$. Hence, if $|\nabla u(z)|^2 \le \exp \mathcal{A}\bigl(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma^{-1}(z))\bigr)$ then

$$P(|\nabla u(z)|) \le \exp \mathcal{A}\bigl(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma^{-1}(z))\bigr). \tag{1-27}$$

If, however, $|\nabla u(z)|^2 > \exp \mathcal{A}\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma^{-1}(z))\big)$, then

$$P(|\nabla u(z)|) = \frac{|\nabla u(z)|^2}{\mathcal{A}^{-1}(\log|\nabla u(z)|^2)} \le \frac{|\nabla u(z)|^2}{\operatorname{tr}(\sigma^{-1}(z))} \le \nabla u(z)\cdot\sigma(z)\nabla u(z). \tag{1-28}$$

Thus at a.e. $z \in \Omega$, we have

$$P(|\nabla u(z)|) \le \exp \mathcal{A}\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma^{-1}(z))\big) + \nabla u(z)\cdot\sigma(z)\nabla u(z). \tag{1-29}$$

Similar arguments give pointwise bounds for $P(|\sigma(z)\nabla u(z)|)$. Summing these estimates and integrating these pointwise estimates over Ω proves the claim. □

In the following, we say that $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$ is in the Orlicz space $W^{1,P}(\Omega)$ if

$$\int_\Omega P(|\nabla u(z)|)\, dm(z) < \infty.$$

There are further important reasons that make the gauge $P(t)$ a natural and useful choice. For instance, in constructing a minimizer for the energy $A_\sigma[u]$, we are faced with the problem of possible equicontinuity of Sobolev functions with $A_\sigma[u]$ uniformly bounded. In view of Lemma 1.10, this is reduced to describing those weight functions $\mathcal{A}(t)$ for which the condition $P(|\nabla u(z)|) \in L^1(\Omega)$ implies that the continuity modulus of u can be estimated. As we will see later in (3-14), this follows for weakly monotone functions u (in particular, for homeomorphisms), as soon as the divergence condition

$$\int_1^\infty \frac{P(t)}{t^3}\, dt = \infty \tag{1-30}$$

is satisfied; that is, $P(t)$ has almost quadratic growth. In fact, note that the divergence of the integral $\int_1^\infty (\mathcal{A}(t)/t^2)\, dt$ is equivalent to

$$\int_1^\infty \frac{P(t)}{t^3}\, dt = \frac{1}{2}\int_1^\infty \frac{\mathcal{A}'(t)}{t}\, dt = \frac{1}{2}\int_1^\infty \frac{\mathcal{A}(t)}{t^2}\, dt = \infty, \tag{1-31}$$

where we have used the substitution $\mathcal{A}(s) = \log(t^2)$. Thus the condition (1-25) is directly connected to the smoothness properties of solutions of finite energy for conductivities satisfying (1-24).

We are now ready to formulate our third main theorem for uniqueness for the inverse problem, which gives a sharp result for singular anisotropic conductivities with a determinant bounded from above and below by positive constants.

Theorem 1.11. *Let $\Omega \subset \mathbb{C}$ be a bounded simply connected domain with smooth boundary and $\mathcal{A} : [1,\infty) \to [0,\infty)$ be a strictly increasing smooth function satisfying the almost linear growth condition* (1-25). *Let $\sigma_1, \sigma_2 \in \Sigma(\Omega)$ be matrix-valued conductivities in Ω which satisfy the integrability condition*

$$\int_\Omega \exp\big(\mathcal{A}\big(\operatorname{tr}\sigma(z) + \operatorname{tr}(\sigma(z)^{-1})\big)\big)\, dm(z) < \infty. \tag{1-32}$$

Moreover, suppose that $c_1 \le \det(\sigma_j(z)) \le c_2$, with $z \in \Omega$, $j = 1, 2$, for some $c_1, c_2 > 0$, and $L_{\sigma_1} = L_{\sigma_2}$. Then there is a $W^{1,1}_{\mathrm{loc}}$-homeomorphism $F : \Omega \to \Omega$ satisfying $F|_{\partial\Omega} = \mathrm{id}$ such that

$$\sigma_1 = F_*\sigma_2.$$

We note that the determination of σ from L_σ in Theorems 1.8, 1.9, and 1.11 is constructive in the sense that one can write an algorithm which constructs σ from Λ_σ. For example, for nondegenerate scalar conductivities, such a construction has been numerically implemented in [Astala et al. 2011].

Let us next discuss the borderline of the visibility somewhat formally. Below we say that a conductivity is *visible* if there is an algorithm which reconstructs the conductivity σ from the boundary measurements L_σ, possibly up to a change of coordinates. In other words, for visible conductivities, one can use the boundary measurements to produce an image of the conductivity in the interior of Ω in some deformed coordinates. For simplicity, let us consider conductivities with $\det\sigma$ bounded from above and below. Then, Theorems 1.7 and 1.11 can be interpreted by saying that the almost linear growth condition (1-25) for the weight function $\mathcal{A}$ gives the *borderline of visibility* for the trace of the conductivity matrix: If $\mathcal{A}$ satisfies (1-25), the conductivities satisfying the integrability condition (1-32) are visible. However, if $\mathcal{A}$ does not satisfy (1-25), we can construct a conductivity in Ω satisfying the integrability condition (1-32) which appears as if an obstacle (which does not exist in reality) would have been included in the domain.

Thus the borderline of the visibility is between any spaces $\Sigma_{\mathcal{A}_1}$ and $\Sigma_{\mathcal{A}_2}$, where $\mathcal{A}_1$ satisfies condition (1-25) and $\mathcal{A}_2$ does not satisfy it. Examples of such gauge functions are $\mathcal{A}_1(t) = t(1+\log t)^{-1}$ and $\mathcal{A}_2(t) = t(1+\log t)^{-1-\varepsilon}$ with $\varepsilon > 0$.

Summarizing the results, in terms of the trace of the conductivity, we have identified the borderline of visible conductivities and the borderline of invisibility cloaking conductivities. Moreover, these borderlines are not the same and between the visible and the invisibility cloaking conductivities, there are conductivities creating electric holograms.

2. Proofs for the existence and uniqueness of the solution of the direct problem and for the counterexamples

First we show that under the conditions (1-24) and (1-25), the Dirichlet problem for the conductivity equation admits a unique solution u with finite energy $A_\sigma[u]$.

2A. ***The Dirichlet problem.*** In this section we prove Theorem 1.4. In fact, we prove it in a more general setting than it was stated.

Theorem 2.1. *Let $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$, where $\mathcal{A}$ satisfies the almost linear growth condition* (1-25). *Then, if $h \in H^{1/2}(\partial\Omega)$ is such that $L_\sigma[h] < \infty$ and $X = \{v \in W^{1,1}(\Omega) : v|_{\partial\Omega} = h\}$, there is a unique $w \in X$ satisfying* (1-10). *Moreover, w satisfies the conductivity equation*

$$\nabla\cdot\sigma\nabla w = 0 \quad \text{in } \Omega \tag{2-1}$$

in the sense of distributions, and has the regularity $w \in W^{1,P}(\Omega)$.

Proof. For $N > 0$, define $\Omega_N = \{x \in \Omega : \|\sigma(x)\| + \|\sigma(x)^{-1}\| \le N\}$. Let $w_n \in X$ be such that

$$\lim_{n\to\infty} A_\sigma[w_n] = C_0 = \inf\{A_\sigma[v] : v \in X\} = L_\sigma[h] < \infty$$

and $A_\sigma[w_n] < C_0 + 1$. Then by Lemma 1.10,

$$\int_\Omega P(|\nabla w_n(x)|)\,dm(x) + \int_\Omega P(|\sigma(x)\nabla w_n(x)|)\,dm(x) \le 2(C_1 + C_0 + 1) = C_2, \tag{2-2}$$

where

$$C_1 = \int_\Omega e^{\mathcal{A}(K(z))}\, dm(z).$$

By [Astala et al. 2009, Lemmas 20.5.3, 20.5.4], there is a convex and unbounded function $\Phi : [0, \infty) \to \mathbb{R}$ such that

$$\Phi(t) \le P(t) + c_0 \le 2\Phi(t)$$

with some $c_0 > 0$ and moreover, the function $t \mapsto \Phi(t^{5/8})$ is convex and increasing. This implies that $P(t) \ge c_1 t^{8/5} - c_2$ for some $c_1 > 0$, $c_2 \in \mathbb{R}$. Thus (2-2) yields that for all $1 < q \le 8/5$,

$$\|\nabla w_n\|_{L^q(\Omega)} \le C_3 = C_3(q, C_0, C_1) \quad \text{for } n \in \mathbb{Z}_+.$$

Using the Poincaré inequality in $L^q(\Omega)$ and that $(w_n - w_1)|_{\partial\Omega} = 0$, we see that

$$\|w_n - w_1\|_{L^q(\Omega)} \le C_4 C_3.$$

Thus, there is C_5 such that $\|w_n\|_{W^{1,q}(\Omega)} < C_5$ for all n. By restricting to a subsequence of $(w_n)_{n=1}^\infty$, which we denote in the sequel also by w_n, we see, using the Banach–Alaoglu theorem, that w_n converges as $n \to \infty$ to a limit in $W^{1,q}(\Omega)$. We denote this limit by w. As $W^{1,q}(\Omega)$ embeds compactly to $H^s(\Omega)$ for $s < 2(1 - q^{-1})$, we see that $\|w_n - w\|_{H^s(\Omega)} \to 0$ as $n \to \infty$ for all $s \in \left(\frac{1}{2}, \frac{3}{4}\right)$. Thus $w_n|_{\partial\Omega} \to w|_{\partial\Omega}$ in $H^{s-1/2}(\partial\Omega)$ as $n \to \infty$. This implies that $w|_{\partial\Omega} = h$ and $w \in X$. Moreover, for any $N > 0$,

$$\frac{1}{N}\int_{\Omega_N} |\nabla w_n(x)|^2\, dm(x) \le \int_{\Omega_N} \nabla w_n(x) \cdot \sigma(x) \nabla w_n(x)\, dm(x) \le C_0 + 1.$$

This implies that $\nabla w_n|_{\Omega_N}$ are uniformly bounded in $L^2(\Omega_N)$. Thus by restricting to a subsequence, we can assume that $\nabla w_n|_{\Omega_N}$ converges weakly in $L^2(\Omega_N)^2$ as $n \to \infty$. Clearly, the weak limit must be $\nabla w|_{\Omega_N}$. Since the norm

$$V \mapsto \left(\int_{\Omega_N} V \cdot \sigma V\, dm\right)^{1/2}$$

in $L^2(\Omega_N)^2$ is weakly lower semicontinuous, we see that

$$\int_{\Omega_N} \nabla w(x) \cdot \sigma(x) \nabla w(x)\, dm(x) \le \liminf_{n\to\infty} \int_{\Omega_N} \nabla w_n(x) \cdot \sigma(x) \nabla w_n(x)\, dm(x) \le C_0.$$

As this holds for all N, we see by applying the monotone convergence theorem as $N \to \infty$ that (1-10) holds. Thus w is a minimizer of A_σ in X.

By the above, $\sigma \nabla w_n \to \sigma \nabla w$ weakly in $L^2(\Omega_N)$ as $n \to \infty$ for all N. As noted above, there is a convex function $\Phi : [0, \infty) \to \mathbb{R}$ such that

$$\Phi(t) \le P(t) + c_0 \le 2\Phi(t), \quad c_0 > 0,$$

and $\Phi(t)$ is increasing for large values of t. Thus it follows from the semicontinuity results for integral operators, [Attouch et al. 2006, Theorem 13.1.2], Lebesgue's monotone convergence theorem, and (2-2) that

$$\begin{aligned}\int_\Omega \big(\Phi(|\nabla w|)+\Phi(|\sigma\nabla w|)\big)\,dm(x) &\le \lim_{N\to\infty}\liminf_{n\to\infty}\int_{\Omega_N}\big(\Phi(|\nabla w_n|)+\Phi(|\sigma\nabla w_n|)\big)\,dm\\ &\le \lim_{N\to\infty}\liminf_{n\to\infty}\int_{\Omega_N}\big(P(|\nabla w_n|)+P(|\sigma\nabla w_n|)\big)\,dm+2c_0|\Omega|\\ &\le C_2+2c_0|\Omega|.\end{aligned}$$

It follows from the above and the inequality $P(t)\ge c_1t^{8/5}-c_2$ that $\sigma(x)\nabla w(x)\in L^1(\Omega)$. Consider next $\phi\in C_0^\infty(\Omega)$. As $w+t\phi\in X$, $t\in\mathbb{R}$, and as w is a minimizer of A_σ in X, it follows that

$$\frac{d}{dt}A_\sigma[w+t\phi]\Big|_{t=0}=2\int_\Omega \nabla\phi(x)\cdot\sigma(x)\nabla w(x)\,dm(x)=0.$$

This shows that the conductivity equation (2-1) is valid in the sense of distributions.

Next, assume that w and $\tilde w$ are both minimizers of A_σ in X. Using the convexity of A_σ, we see that then the second derivative of $t\mapsto A_\sigma[tw+(1-t)\tilde w]$ vanishes at $t=0$. This implies that $\nabla(w-\tilde w)=0$ for a.e. $x\in\Omega$. As w and $\tilde w$ coincide at the boundary, this yields that $w=\tilde w$ and thus the minimizer is unique. □

The fact that the minimizer w is continuous will be proven in the next subsection.

2B. ***The Beltrami equation.*** It is natural to ask if the minimizer w in (1-10) is the only solution of finite σ-energy $A_\sigma[w]$ to the boundary value problem

$$\begin{aligned}\nabla\cdot\sigma\nabla w&=0\quad\text{in }\Omega,\\ w|_{\partial\Omega}&=h.\end{aligned}\tag{2-3}$$

It turns out that this is the case and to prove this we introduce one of the basic tools in this work, the Beltrami differential equation.

To this end, recall the Hodge-star operator $*$, which in two dimensions is just the rotation

$$*=\begin{pmatrix}0&-1\\1&0\end{pmatrix}.$$

Note that $\nabla\cdot(*\nabla w)=0$ for all $w\in W^{1,1}(\Omega)$ and recall that $\Omega\subset\mathbb{C}$ is simply connected. If $\sigma(x)=[\sigma^{jk}(x)]_{j,k=1}^2\in\Sigma_{\mathcal{A}}(\Omega)$, where $\mathcal{A}$ satisfies (1-19), and if $u\in W^{1,1}(\Omega)$ is a distributional solution to the conductivity equation

$$\nabla\cdot\sigma(x)\nabla u(x)=0,\tag{2-4}$$

then by Lemma 1.10, we have $P(\nabla u)$, $P(\sigma\nabla u)\in L^1(\Omega)$ and thus in particular $\sigma\nabla u\in L^1(\Omega)$. By (2-4) and the Poincaré lemma, there is a function $v\in W^{1,1}(\Omega)$ such that

$$\nabla v=*\,\sigma(x)\nabla u(x).\tag{2-5}$$

Then

$$\nabla\cdot\sigma^*(x)\nabla v=0\quad\text{in }\Omega,\qquad \sigma^*(x)=*\,\sigma(x)^{-1}*^t.\tag{2-6}$$

In particular, the above shows that $u,v\in W^{1,P}(\Omega)$. Moreover, an explicit calculation (see, e.g., [Astala et al. 2009, Formula (16.20)]) reveals that the function $f=u+iv$ satisfies

$$\partial_{\bar z}f=\mu\,\partial_z f+\nu\,\overline{\partial_z f},\tag{2-7}$$

where

$$\mu = \frac{\sigma^{22} - \sigma^{11} - 2i\sigma^{12}}{1 + \mathrm{tr}(\sigma) + \det(\sigma)}, \qquad \nu = \frac{1 - \det(\sigma)}{1 + \mathrm{tr}(\sigma) + \det(\sigma)}, \tag{2-8}$$

and $\partial_{\bar z} = \frac{1}{2}(\partial_{x_1} + i\,\partial_{x_2})$ with $\partial_z = \frac{1}{2}(\partial_{x_1} - i\,\partial_{x_2})$. Note that $|\mu(z)| + |\nu(z)| < 1$ for a.e. z. Summarizing, for $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$, any distributional solution $u \in W^{1,1}(\Omega)$ of (2-4) is the real part of a solution f of (2-7). Conversely, the real part of any solution $f \in W^{1,1}(\Omega)$ of (2-7) satisfies (2-4), while the imaginary part is a solution to (2-6) and as $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$, (2-4)–(2-6) and Lemma 1.10 yield that $u, v \in W^{1,P}(\Omega)$, and hence $f \in W^{1,P}(\Omega)$.

Furthermore, the ellipticity bound of $\sigma(z)$ is closely related to the distortion of the mapping f. Indeed, in the case when $\sigma(z_0) = \mathrm{diag}(\lambda_1, \lambda_2)$, a direct computation shows that

$$K_\sigma(z_0) = K_{\mu,\nu}(z_0), \quad \text{where } K_{\mu,\nu}(z) = \frac{1 + |\mu(z)| + |\nu(z)|}{1 - (|\mu(z)| + |\nu(z)|)} \tag{2-9}$$

and $K_\sigma(z)$ is the ellipticity of $\sigma(z)$ defined in (1-5). Using the chain rule for the complex derivatives, which can be written as

$$\partial(v \circ F) = (\partial v) \circ F \cdot \partial F + (\bar\partial v) \circ F \cdot \overline{\bar\partial F}, \tag{2-10}$$

$$\bar\partial(v \circ F) = (\partial v) \circ F \cdot \bar\partial F + (\bar\partial v) \circ F \cdot \overline{\partial F}, \tag{2-11}$$

we see that $|\mu(z)|$ and $|\nu(z)|$ do not change in an orthogonal rotation of the coordinate axis, $z \mapsto \alpha z$, where $\alpha \in \mathbb{C}$, $|\alpha| = 1$. Since, for any $z_0 \in \Omega$ there exists an orthogonal rotation of the coordinate axis so that matrix $\sigma(z_0)$ is diagonal in the rotated coordinates, we see that the identity (2-9) holds for all $z_0 \in \Omega$.

Equation (2-7) is also equivalent to the Beltrami equation

$$\bar\partial f(z) = \tilde\mu(z)\,\partial f(z) \quad \text{in } \Omega, \tag{2-12}$$

with the Beltrami coefficient

$$\tilde\mu(z) = \begin{cases} \mu(z) + \nu(z)\,\partial_z f(x)(\overline{\partial_z f(x)})^{-1} & \text{if } \partial_z f(x) \neq 0, \\ \mu(z) & \text{if } \partial_z f(x) = 0 \end{cases} \tag{2-13}$$

satisfying $|\tilde\mu(z)| \le |\mu(z)| + |\nu(z)|$ pointwise. We define the distortion of f at z to be

$$K(z, f) := K_{\tilde\mu}(z) = \frac{1 + |\tilde\mu(z)|}{1 - |\tilde\mu(z)|} \le K_\sigma(z), \quad z \in \Omega. \tag{2-14}$$

Below we will also use the notation $K(z, f) = K_f(z)$.

In the sequel we will use frequently these different interpretations of the Beltrami equation. Note that

$$K(z, f) = \frac{1 + |\tilde\mu(z)|}{1 - |\tilde\mu(z)|}$$

so that

$$K(z, f) = \frac{|\partial f| + |\bar\partial f|}{|\partial f| - |\bar\partial f|}.$$

As $\|Df\|^2 = (|\partial f| + |\bar\partial f|)^2$ and $J(z, f) = |\partial f|^2 - |\bar\partial f|^2$, this yields the distortion equality (see, e.g., [Astala et al. 2009, Formula (20.3)])

$$\|Df(z)\|^2 = K(z, f)\, J(z, f) \quad \text{for a.e. } z \in \Omega. \tag{2-15}$$

We will use extensively the fact that if a homeomorphism $F : \Omega \to \Omega'$, $F \in W^{1,1}(\Omega)$, is a finite distortion mapping with the distortion $K_F \in L^1(\Omega)$ then by [Hencl et al. 2005] or [Astala et al. 2009, Theorem 21.1.4] the inverse function $H = F^{-1} : \Omega' \to \Omega$ is in $W^{1,2}(\Omega')$ and its derivative DH satisfies

$$\|DH\|_{L^2(\Omega')} \le 2\|K_F\|_{L^1(\Omega)}. \tag{2-16}$$

We will also need a few basic notions (see [Astala et al. 2009]) from the theory of Beltrami equations. As the coefficients μ, ν are defined only in the bounded domain Ω, outside Ω we set $\mu(z) = \nu(z) = 0$ and $\sigma(z) = 1$, and consider global solutions to (2-7) in $\mathbb{C}$. In particular, we consider the case when Ω is the unit disc $\mathbb{D} = B(1)$. We say that a solution $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{C})$ of (2-7) in $z \in \mathbb{C}$ is a *principal solution* if

(1) $f : \mathbb{C} \to \mathbb{C}$ is a homeomorphism of $\mathbb{C}$ and

(2) $f(z) = z + \mathcal{O}(1/z)$ as $z \to \infty$.

The existence of principal solutions is a fundamental fact that holds true in quite wide generality. Further, with the principal solution one can classify all solutions, of sufficient regularity, to the Beltrami equation. These facts are summarized in the following version of Stoilow's factorization theorem (see [Astala et al. 2009, Theorem 20.5.2] for the proof).

Theorem 2.2. *Suppose $\mu(z)$ is supported in the unit disk $\mathbb{D}$, $|\mu(z)| < 1$ a.e. and*

$$\int_{\mathbb{D}} \exp(\mathcal{A}(K_\mu(z)))\, dm(z) < \infty, \qquad K_\mu(z) = \frac{1 + |\mu(z)|}{1 - |\mu(z)|},$$

where $\mathcal{A}$ satisfies the almost linear growth condition (1-25)*. Then the equation*

$$\bar\partial\Phi(z) = \mu(z)\, \partial\Phi(z), \quad z \in \mathbb{C}, \tag{2-17}$$

$$\Phi(z) = z + \mathcal{O}(1/z) \qquad \text{as } z \to \infty, \tag{2-18}$$

has a unique solution in $\Phi \in W^{1,1}_{\mathrm{loc}}(\mathbb{C})$. The solution $\Phi : \mathbb{C} \to \mathbb{C}$ is a homeomorphism and satisfies $\Phi \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$. Moreover, when $\Omega_1 \subset \mathbb{C}$ is open, every solution of the equation

$$\bar\partial f(z) = \mu(z)\, \partial f(z), \quad z \in \Omega_1, \tag{2-19}$$

with the regularity $f \in W^{1,P}_{\mathrm{loc}}(\Omega_1)$, can be written as $f = H \circ \Phi$, where Φ is the solution to (2-17)–(2-18) *and H is a holomorphic function in $\Omega_1' = \Phi(\Omega_1)$.*

Below we combine this result with the Poincaré lemma to analyze the solutions of the conductivity equation in the simply connected domain Ω.

Corollary 2.3. *Let $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$, where $\mathcal{A}$ satisfies* (1-25)*, and $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$ satisfy*

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{in } \Omega \quad \text{and} \quad \int_\Omega \nabla u(x) \cdot \sigma(x) \nabla u(x)\, dm(x) < \infty. \tag{2-20}$$

Then there exists a homeomorphism $\Phi : \mathbb{C} \to \mathbb{C}$, $\Phi \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$, *and a harmonic function* w, *defined in the domain* $\Omega' = \Phi(\Omega)$, *such that* $u = w \circ \Phi$. *In particular*, $u : \Omega \to \mathbb{R}$ *is continuous.*

Proof. Let $v \in W^{1,1}_{\mathrm{loc}}(\Omega)$ be the conjugate function of u, described in (2-5), and set $f = u + iv$. Then by Lemma 1.10, we have $f \in W^{1,P}(\Omega)$, and Theorem 2.2 yields that $f = H \circ \Phi$, where $\Phi : \mathbb{C} \to \mathbb{C}$ is a homeomorphism with $\Phi \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$ and H is holomorphic in $\Phi(\Omega)$. Thus the real part $u = (\mathrm{Re}\, H) \circ \Phi$ has the required factorization with $w = \mathrm{Re}\, H$. □

Theorem 2.1 and Corollary 2.3 yield Theorem 1.4.

2C. ***Invariance of the Dirichlet-to-Neumann form under coordinate transformations.*** In this section, we assume that $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$, where $\mathcal{A}$ satisfies (1-25). We say that $F : \Omega \to \Omega'$ satisfies the condition $\mathcal{N}$ if for any measurable set $E \subset \Omega$, we have $|E| = 0 \Rightarrow |F(E)| = 0$. Also, we say that F satisfies the condition $\mathcal{N}^{-1}$ if for any measurable set $E \subset \Omega$, we have $|F(E)| = 0 \Rightarrow |E| = 0$.

Let $\sigma \in \Sigma_{\mathcal{A}}(\mathbb{C})$ be such that σ is constantly 1 in $\mathbb{C} \setminus \Omega$. Let

$$\hat{\mu}(z) = \frac{\sigma^{11}(z) - \sigma^{22}(z) + 2i\sigma^{12}(z)}{\sigma^{11}(z) + \sigma^{22}(z) + 2\sqrt{\det \sigma(z)}} \tag{2-21}$$

be the Beltrami coefficient associated to the isothermal coordinates corresponding to σ; see, e.g., [Sylvester 1990; Astala et al. 2009, Theorem 10.1.1]. A direct computation shows that $K_{\hat{\mu}}(z) = K_\sigma(z)$ and thus

$$\exp(\mathcal{A}(K_{\hat{\mu}})) \in L^1_{\mathrm{loc}}(\mathbb{C}),$$

and by Theorem 2.2, there exists a homeomorphism $F : \mathbb{C} \to \mathbb{C}$ satisfying (2-17)–(2-18) with the Beltrami coefficient $\hat{\mu}$ such that $F \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$. Due to the choice of $\hat{\mu}$, the conductivity $F_*\sigma$ is isotropic; see, e.g., [Sylvester 1990; Astala et al. 2009, Theorem 10.1.1]. Let us next consider the properties of the map F. First, as

$$\exp(\mathcal{A}(K_{\hat{\mu}})) \in L^1_{\mathrm{loc}}(\mathbb{C}),$$

it follows from [Kauhanen et al. 2003] that the function F satisfies the condition $\mathcal{N}$. Moreover, the fact that $K_F = K_{\hat{\mu}} \in L^1_{\mathrm{loc}}(\mathbb{C})$ implies by (2-16) that its inverse $H = F^{-1}$ is in $W^{1,2}_{\mathrm{loc}}(\mathbb{C})$. This yields by [Astala et al. 2009, Theorem 3.3.7] that F^{-1} satisfies the condition $\mathcal{N}$. In particular, the above yields that both F and F^{-1} are in $W^{1,P}_{\mathrm{loc}}(\mathbb{C})$.

The following lemma formulates the invariance of the Dirichlet-to-Neumann forms in the diffeomorphisms satisfying the above properties.

Lemma 2.4. *Assume that* $\Omega, \widetilde{\Omega} \subset \mathbb{C}$ *are bounded*, *simply connected domains with smooth boundaries and that* $\sigma \in \Sigma_{\mathcal{A}}(\Omega)$ *and* $\tilde{\sigma} \in \Sigma_{\mathcal{A}}(\widetilde{\Omega})$, *where* $\mathcal{A}$ *satisfies* (1-25). *Let* $F : \Omega \to \widetilde{\Omega}$ *be a homeomorphism so that* F *and* F^{-1} *are* $W^{1,P}$*-smooth and* F *satisfies conditions* $\mathcal{N}$ *and* $\mathcal{N}^{-1}$. *Assume that* F *and* F^{-1} *are* C^1*-smooth near the boundary and assume that* $\rho = F|_{\partial\Omega}$ *is* C^2*-smooth. Also*, *suppose* $\tilde{\sigma} = F_*\sigma$. *Then*

$$L_{\tilde{\sigma}}[\tilde{h}] = L_\sigma[\tilde{h} \circ \rho]$$

for all $\tilde{h} \in H^{1/2}(\partial\widetilde{\Omega})$.

Proof. As F has the properties $\mathcal{N}$ and $\mathcal{N}^{-1}$, we have the area formula

$$\int_{\widetilde{\Omega}} H(y)\,dm(y) = \int_{\Omega} H(F(x))\,J(x,F)\,dm(x) \tag{2-22}$$

for all simple functions $H:\widetilde{\Omega}\to\mathbb{C}$, where $J(x,F)$ is the Jacobian determinant of F at x. Thus (2-22) holds for all $H\in L^1(\widetilde{\Omega})$.

Let $\tilde h\in H^{1/2}(\partial\widetilde{\Omega})$ and assume that $L_{\tilde\sigma}[\tilde h]<\infty$. Let $\tilde u:\widetilde{\Omega}\to\mathbb{R}$ be the unique minimizer of $A_{\tilde\sigma}[v]$ in $\widetilde{X}=\{\tilde v\in W^{1,1}(\widetilde{\Omega}):\tilde v|_{\partial\widetilde{\Omega}}=\tilde h\}$. Then $\tilde u$ is the solution of the conductivity equation

$$\nabla\cdot\tilde\sigma\nabla\tilde u=0,\quad \tilde u|_{\partial\widetilde{\Omega}}=\tilde h. \tag{2-23}$$

We define $h=\tilde h\circ F|_{\partial\Omega}$ and $u=\tilde u\circ F:\Omega\to\mathbb{C}$.

By Corollary 2.3, $\tilde u$ can be written in the form $\tilde u=\tilde w\circ\widetilde{G}$, where $\tilde w$ is harmonic and $\widetilde{G}\in W^{1,1}_{\text{loc}}(\mathbb{C})$ is a homeomorphism $\widetilde{G}:\mathbb{C}\to\mathbb{C}$.

By the Gehring–Lehto theorem (see [Astala et al. 2009, Corollary 3.3.3]), a homeomorphism $F\in W^{1,1}_{\text{loc}}(\Omega)$ is differentiable almost everywhere in Ω, say in the set $\Omega\setminus A$, where A has Lebesgue measure zero. Similar arguments for $\widetilde{G}$ show that $\widetilde{G}$ and the solution $\tilde u$ are differentiable almost everywhere, say in the set $\widetilde{\Omega}\setminus A'$, where A' has Lebesgue measure zero.

Since F has the property $\mathcal{N}^{-1}$, we see that $A''=A'\cup F^{-1}(A')\subset\Omega$ has measure zero, and for $x\in\Omega\setminus A''$, the chain rule gives

$$Du(x)=(D\tilde u)(F(x))\cdot DF(x). \tag{2-24}$$

Note that the facts that F is a map with an exponentially integrable distortion and that $\tilde u$ is a real part of a map with an exponentially integrable distortion, do not generally imply, at least according to the knowledge of the authors, that their composition u is in $W^{1,1}_{\text{loc}}(\Omega)$. To overcome this problem, we define for $m>1$,

$$\widetilde{\Omega}_m=\bigl\{y\in\widetilde{\Omega}:\|DF^{-1}(y)\|+\|DF(F^{-1}(y))\|+\|\tilde\sigma(y)\|+|\nabla\tilde u(y)|<m\bigr\}$$

and $\Omega_m=F^{-1}(\widetilde{\Omega}_m)$. Then $\nabla u|_{\Omega_m}\in L^2(\Omega_m)$ and $\|\sigma\|<m^5$ in Ω_m; see (1-13).

Now for any $m>0$,

$$\int_{\widetilde{\Omega}_m}\nabla\tilde u(y)\cdot\tilde\sigma(y)\nabla\tilde u(y)\,dm(y)\le A_{\tilde\sigma}[\tilde u]<\infty. \tag{2-25}$$

Due to the definition of $\tilde\sigma=F_*\sigma$, we see by using formulae (2-22) and (2-24) that

$$\int_{\Omega_m}\nabla u(x)\cdot\sigma(x)\nabla u(x)\,dm(x)=\int_{\widetilde{\Omega}_m}\nabla\tilde u(y)\cdot\tilde\sigma(y)\nabla\tilde u(y)\,dm(y). \tag{2-26}$$

Letting $m\to\infty$ and using the monotone convergence theorem, we see that

$$\int_{\Omega}\nabla u(x)\cdot\sigma(x)\nabla u(x)\,dm(x)=\int_{\widetilde{\Omega}}\nabla\tilde u(y)\cdot\tilde\sigma(y)\nabla\tilde u(y)\,dm(y)=A_{\tilde\sigma}[\tilde u]<\infty. \tag{2-27}$$

By Lemma 1.10, this implies that $u\in W^{1,P}(\Omega)\subset W^{1,1}(\Omega)$.

Clearly, as $\rho=F|_{\partial\Omega}$ is C^2-smooth, $h:=\tilde h\circ F\in H^{1/2}(\partial\Omega)$ and $u|_{\partial\Omega}=h$. Thus

$$u\in X=\{w\in W^{1,1}(\Omega):w|_{\partial\Omega}=h\}.$$

Since $\tilde u$ is a minimizer of $A_{\tilde\sigma}$ in $\widetilde X$, and u satisfies

$$A_\sigma[u] \le A_{\tilde\sigma}[\tilde u] = L_{\tilde\sigma}(\tilde h),$$

we see that

$$L_\sigma[h] \le L_{\tilde\sigma}[\tilde h].$$

Changing the roles of $\tilde\sigma$ and σ, we obtain an opposite inequality, and prove the claim. □

In particular, if $\sigma \in \Sigma_{\mathcal A}(\Omega)$, $\tilde\sigma \in \Sigma_{\mathcal A}(\widetilde\Omega)$ and F are as in Lemma 2.4 and in addition to that, σ and $\tilde\sigma$ are bounded near $\partial\Omega$ and $\partial\widetilde\Omega$ respectively and $\rho = F|_{\partial\Omega} : \partial\Omega \to \partial\widetilde\Omega$ is C^2-smooth, then the quadratic forms L_σ and $L_{\tilde\sigma}$ can be written in terms of the Dirichlet-to-Neumann maps $\Lambda_\sigma : H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)$ and $\Lambda_{\tilde\sigma} : H^{1/2}(\partial\widetilde\Omega) \to H^{-1/2}(\partial\tilde\Omega)$ as in formula (1-4). Then, Lemma 2.4 implies that

$$\Lambda_{\tilde\sigma} = \rho_* \Lambda_\sigma, \tag{2-28}$$

where $\rho_*\Lambda_\sigma$ is the push-forward of Λ_σ in ρ defined by

$$(\rho_*\Lambda_\sigma)(\tilde h) = j \cdot \big((\Lambda_\sigma(\tilde h \circ \rho)) \circ \rho^{-1}\big)$$

for $\tilde h \in H^{1/2}(\partial\tilde\Omega)$, where $j(z)$ is the Jacobian of the map $\rho^{-1} : \partial\widetilde\Omega \to \partial\Omega$.

2D. ***Counterexamples revisited.*** In this section we give the proofs of the claims stated in Section 1B. We start by proving Theorem 1.6. Since the change of variables used in the integration is singular, we present the arguments in detail.

Proof of Theorem 1.6. (i) Our aim is first to show that we have $L_\sigma[h] \le L_{\tilde\sigma}[h]$ and then to prove the opposite inequality. The proofs of these inequalities are based on different techniques due to the fact that $\tilde\sigma$ is not even in $L^1(B(2))$.

Let $0 < r < 2$ and

$$\mathcal K(r) = \mathcal K \cup F(\overline B(r)).$$

Moreover, let $\tilde\sigma_r$ be a conductivity that coincides with $\tilde\sigma$ in $B(2) \setminus \mathcal K(r)$ and is zero in $\mathcal K(r)$. Similarly, let σ_r be a conductivity that coincides with σ in $B(2) \setminus \overline B(r)$ and is zero in $\overline B(r)$. For these conductivities, we define the quadratic forms $A^r : W^{1,1}(B(2)) \to \mathbb R_+ \cup \{0, \infty\}$ and $\tilde A^r : W^{1,1}(B(2)) \to \mathbb R_+ \cup \{0, \infty\}$,

$$A^r[v] = \int_{B(2)\setminus\overline B(r)} \nabla v \cdot \sigma \nabla v \, dm(x), \quad \tilde A^r[v] = \int_{B(2)\setminus\mathcal K(r)} \nabla v \cdot \tilde\sigma \nabla v \, dm(x).$$

If we minimize $\tilde A^r[v]$ over $v \in W^{1,1}(B(2))$ with $v|_{\partial B(2)} = h$, we see that minimizers exist and that the restriction of any minimizer to $B(2) \setminus \overline{\mathcal K}(r)$ is the function $\tilde u_r \in W^{1,2}(B(2) \setminus \mathcal K(r))$ satisfying

$$\nabla \cdot \tilde\sigma \nabla \tilde u_r = 0 \quad \text{in } B(2) \setminus \mathcal K(r), \qquad \tilde u_r|_{\partial B(2)} = h, \qquad \nu \cdot \tilde\sigma \nabla \tilde u_r|_{\partial\mathcal K(r)} = 0.$$

Analogous equations hold for the minimizer u^r of A^r. As σ in $\overline B(2) \setminus B(r)$ and $\tilde\sigma$ in $\overline{B(2) \setminus \mathcal K(r)}$ are bounded from above and below by positive constants, we see using the change of variables and the chain rule that

$$L_{\sigma_r}[h] = L_{\tilde\sigma_r}[h] \quad \text{for } h \in H^{1/2}(\partial B(2)). \tag{2-29}$$

As $\sigma(x) \geq \sigma_r(x)$ and $\tilde\sigma(x) \geq \tilde\sigma_r(x)$ for all $x \in B(2)$,

$$L_\sigma[h] \geq L_{\sigma_r}[h], \quad L_{\tilde\sigma}[h] \geq L_{\tilde\sigma_r}[h]. \tag{2-30}$$

Let us consider the minimization problem (1-3) for σ. It is solved by the unique minimizer $u \in W^{1,1}(B(2))$ satisfying

$$\nabla\cdot\sigma\nabla u = 0 \quad \text{in } B(2), \quad u|_{\partial B(2)} = h.$$

As $\sigma, \sigma^{-1} \in L^\infty(B(2))$, we have $u \in W^{1,2}(B(2))$ and Morrey's theorem [1938] yields that the solution u is $C^{0,\alpha}$-smooth in the open ball $B(2)$ for some $\alpha > 0$. Thus $u|_{B(R)}$ is in the *Royden algebra*

$$\mathcal{R}(B(R)) = C(B(R)) \cap L^\infty(B(R)) \cap W^{1,2}(B(R))$$

for all $R < 2$; see [Astala et al. 2009, p. 77].

By, e.g., [Iwaniec and Martin 2001, p. 443], for any $0 < R < 2$, the p-capacity of the disc $B(r)$ in $B(R)$ goes to zero as $r \to 0$ for all $p > 1$. Using this, and that $u \in W^{1,2}(B(2)) \subset L^q(B(2))$ for $q < \infty$, we see that (see [Kohn et al. 2008] for explicit estimates in the case when $\sigma = 1$)

$$\lim_{r\to 0} L_{\sigma_r}[h] = L_\sigma[h];$$

that is, the effect of an insulating disc of radius r in the boundary measurements vanishes as $r \to 0$. This and the inequalities (2-29) and (2-30) yield $L_{\tilde\sigma}[h] \geq L_\sigma[h]$. Next we consider the opposite inequality.

Let $\tilde u = u \circ F^{-1}$ in $B(2)\setminus\mathcal{K}$. As F is a homeomorphism, we see that if $x \to 0$ then $d(F(x), \mathcal{K}) \to 0$ and vice versa. Thus, as u is continuous at zero, we see that $\tilde u \in C(B(2)\setminus\mathcal{K}^{\text{int}})$ and $\tilde u$ has the constant value $u(0)$ on $\partial\mathcal{K}$. Moreover, as $F^{-1} \in C^1(B(2)\setminus\mathcal{K})$, we have $\|DF^{-1}\| \leq C_0$ in $B(2)\setminus\mathcal{K}$ and u is in the Royden algebra $\mathcal{R}(B(R))$ for all $R < 2$; we have by [Astala et al. 2009, Theorem 3.8.2] that the chain rule holds implying that $D\tilde u = ((Du)\circ F^{-1})\cdot DF^{-1}$ a.e. in $B(2)\setminus\mathcal{K}$. Let $0 < R' < R'' < 2$. Then

$$|D\tilde u(z)| \leq C_0\|Du\|_{C(\bar B(R''))} \quad \text{for } z \in F(B(R''))\setminus\mathcal{K}.$$

As F and F^{-1} are C^1-smooth up to $\partial B(2)$, we have $\tilde u \in W^{1,1}(B(2)\setminus B(R'))$. These give $\tilde u \in W^{1,1}(B(2)\setminus\mathcal{K})$. Let $\tilde v \in W^{1,1}(B(2))$ be a function that coincides with $\tilde u$ in $B(2)\setminus\mathcal{K}$ and with $u(0)$ in $\mathcal{K}$.

Again, using the chain rule and the area formula as in the proof of Lemma 2.4, we see that $\tilde A^r[\tilde v] = A^r[u]$ for $r > 1$. Applying the monotone convergence theorem twice, we obtain

$$L_{\tilde\sigma}[h] \leq A_{\tilde\sigma}[\tilde v] = \lim_{r\to 0}\tilde A^r[\tilde v] = \lim_{r\to 0} A^r[u] = L_\sigma[h]. \tag{2-31}$$

As we have already proven the opposite inequality, this proves the claim (i).

(ii) Assume that $\tilde\sigma$ is a cloaking conductivity obtained by the transformation map F and the background conductivity $\sigma \in L^\infty(B(2))$, $\sigma \geq c_1 > 0$, but that opposite to the claim, we have $\operatorname{tr}(\tilde\sigma) \in L^1(B(2)\setminus\mathcal{K})$. Using formula (1-6) and the facts $\det(\tilde\sigma) = \det(\sigma\circ F^{-1})$ is bounded from above and below by strictly positive constants and $\operatorname{tr}(\tilde\sigma) \in L^1(B(2)\setminus\mathcal{K})$, we see that

$$\operatorname{tr}(\tilde\sigma^{-1}) = \operatorname{tr}(\tilde\sigma)/\det(\tilde\sigma) \in L^1(B(2)\setminus\mathcal{K}).$$

Hence by Lemma 1.2, $K_{\tilde\sigma} \in L^1(B(2)\setminus\mathcal{K})$. Let $G : B(2)\setminus\mathcal{K} \to B(2)\setminus\{0\}$ be the inverse map of F. Using the formulas (1-5), (1-13), and (2-15), we see that

$$\|\tilde\sigma(y)\| = \frac{\|DF(x)\cdot\sigma(x)\cdot DF(x)^t\|}{J(x,F)} \ge \frac{\|DF(x)\|^2}{J(x,F)K_\sigma(x)} = \frac{K_F(x)}{K_\sigma(x)}, \quad x = F^{-1}(y).$$

As $K_G = K_F \circ F^{-1}$ (see [Astala et al. 2009, Formula (2.15)] and $\|\tilde\sigma(y)\| \le K_{\tilde\sigma}(y)$), the above yields $K_G \in L^1(B(2)\setminus\mathcal{K})$. Hence, we see using (2-16) that $F = G^{-1}$ is in $W^{1,2}(B(2)\setminus\{0\})$ and

$$\|DF\|_{L^2(B(2)\setminus\{0\})} \le 2\|K_G\|_{L^1(B(2)\setminus\mathcal{K})}.$$

By the removability of singularities in Sobolev spaces (see [Kilpeläinen et al. 2000]), this implies that $F : B(2)\setminus\{0\} \to B(2)\setminus\mathcal{K}$ can be extended to a function $F^{\text{ext}} : B(2) \to \mathbb{C}$, $F^{\text{ext}} \in W^{1,2}(B(2))$. As the distortion K_F of the map F is finite a.e., the map F^{ext} is also a finite distortion map; see [Astala et al. 2009, Definition 20.0.3]. Thus, as $F^{\text{ext}} \in W^{1,2}_{\text{loc}}(B(2))$, it follows from the continuity theorem of finite distortion maps [Astala et al. 2009, Theorem 20.1.1] or [Manfredi 1994] that $F^{\text{ext}} : B(2) \to \mathbb{C}$ is continuous. Let $y_0 = F(0)$. Then the set $F^{\text{ext}}(\bar B(2)) = (\bar B(2)\setminus\mathcal{K}) \cup \{y_0\}$ is not closed as ∂K contains more that one point and thus it is not compact. This is a contradiction with the fact that F^{ext} is continuous. This proves the claim (ii). □

Next we prove the claim concerning the last counterexample.

Proof of Theorem 1.7. Let us start by reviewing the properties of the Iwaniec–Martin maps. Let $\mathcal{A}_1 : [1,\infty] \to [0,\infty]$ be a strictly increasing positive smooth function with $\mathcal{A}_1(1) = 0$ which satisfies the condition (1-19). Then by [Iwaniec and Martin 2001, Theorem 11.2.1], there exists a $W^{1,1}$-homeomorphism $F : B(2)\setminus\{0\} \to B(2)\setminus\bar B(1)$ with Beltrami coefficient μ satisfying

$$\int_{B(2)\setminus\{0\}} \exp\big(\mathcal{A}_1(K_\mu(z))\big)\,dm(z) < \infty, \quad \text{where } K_\mu(z) := \frac{1+|\mu(z)|}{1-|\mu(z)|}. \tag{2-32}$$

The function F can be obtained using the construction procedure of [Astala et al. 2009, Section 20.3] (see [Iwaniec and Martin 2001, Theorem 11.2.1] for the original construction) as follows: Let $S(t)$ be solution of the equation

$$\mathcal{A}_1(S(t)) = 1 + \log(t^{-1}), \quad 0 < t \le 1. \tag{2-33}$$

Then $S : (0,1] \to [1,\infty)$ is a well-defined decreasing function, $S(1) = 1$ and with suitably chosen $c_1 > 0$, the function

$$F(z) = \frac{z}{|z|}\rho(|z|), \quad \rho(s) = 1 + c_1\left(\exp\left(\int_0^s \frac{dt}{tS(t)}\right) - 1\right), \tag{2-34}$$

is a homeomorphism $F : B(2)\setminus\{0\} \to B(2)\setminus\bar B(1)$. We say that F is the Iwaniec–Martin map corresponding to the weight function $\mathcal{A}_1(t)$.

Next let $\mathcal{A} : [1,\infty] \to [0,\infty]$ be a strictly increasing positive smooth function with $\mathcal{A}(1) = 0$ which satisfies the condition (1-19) and let F_1 be the Iwaniec–Martin map corresponding to the weight function $\mathcal{A}_1(t) = \mathcal{A}(4t)$.

Using the inverse of the map F_1, we define $\sigma_1 = (F_1^{-1})_* 1$ on $B(2) \setminus \{0\}$ and consider this function as an a.e. defined measurable function on $B(2)$. Using the definition of push-forward, (2-32), we see that $\det(\sigma_1) = 1$ and

$$K_{\sigma_1}(z) = K(F_1^{-1}(z), F_1^{-1}) = K_\mu(z).$$

Thus Lemma 1.2 and the fact that F_1 satisfies (2-32) with the weight function $\mathcal{A}_1(t) = \mathcal{A}(4t)$ yield that σ_1 satisfies (1-20) with the weight function $\mathcal{A}(t)$.

Recall that the conductivity γ_1 is identically 1 in $B(2) \setminus \bar{B}(1)$ and zero in $\bar{B}(1)$. Next, we consider the minimization problem (1-3) with the conductivities γ_1 and σ_1. To this end, we make analogous definitions to the proof of Theorem 1.6. For $1 < r < 2$, let γ_r be a conductivity that is 1 in $B(2) \setminus B(r)$ and is zero in $B(r)$. Similarly, let σ_r be a conductivity that coincides with σ_1 in $B(2) \setminus B(r-1)$ and is zero in $B(r-1)$.

As in (2-29) and (2-30), we see for $h \in H^{1/2}(\partial B(2))$ and $r > 1$, that

$$L_{\sigma_r}[h] = L_{\gamma_r}[h], \quad L_{\sigma_r}[h] \le L_{\sigma_1}[h], \quad L_{\gamma_r}[h] \le L_{\gamma_1}[h]. \tag{2-35}$$

Let $h \in H^{1/2}(\partial B(2))$. For $1 \le r < 2$, the solution of the boundary value problem

$$\Delta w_r = 0 \quad \text{in } B(2) \setminus \bar{B}(r), \qquad w_r|_{\partial B(2)} = h, \qquad \partial_\nu w_r|_{\partial B(r)} = 0$$

satisfies $L_{\gamma_r}[h] = \|\nabla w_r\|^2_{L^2(B(2)\setminus \bar{B}(r))}$ and it is easy to see that

$$\lim_{r \to 0} L_{\gamma_r}[h] = L_{\gamma_1}[h] \quad \text{for } h \in H^{1/2}(\partial B(2)). \tag{2-36}$$

Let $w = w_1$. Note that $w \in W^{1,2}(B(2) \setminus \bar{B}(1))$.

Let us consider the function $v = w \circ F_1$. As F_1 is C^1-smooth in $\bar{B}(2) \setminus \{0\}$ and the function w is C^1-smooth in $\bar{B}(R) \setminus \bar{B}(1)$ for all $1 < R < 2$, we have by the chain rule that

$$Dv(x) = (Dw)(F_1(x)) \cdot DF_1(x)$$

for all $x \in B(2) \setminus \{0\}$. As $Dw \in L^2(B(2) \setminus B(R))$ and $Dw \in L^\infty(B(R) \setminus \bar{B}(1))$ for all $1 < R < 2$, and

$$DF_1(x) = \frac{\rho(|x|)}{|x|}(I - P(x)) + \rho'(|x|)P(x),$$

where

$$P(x) : y \mapsto |x|^{-2}(x \cdot y)x$$

is the projector to the radial direction at the point x, using (2-34) we see that $\|DF_1(x)\| \le C|x|^{-1}$ with some $C > 0$ and

$$Dv \in L^p(B(2) \setminus \{0\}) \quad \text{for any } p \in (1, 2). \tag{2-37}$$

Thus $v \in W^{1,p}(B(2)\setminus\{0\})$ with any $p \in (1, 2)$ and by the removability of singularities in Sobolev spaces (see, e.g., [Kilpeläinen et al. 2000, Theorem 4.6 and p. 241]), the function v can be considered as a measurable function in $B(2)$ for which $v \in W^{1,p}(B(2))$. Thus v is in the domain of definition of the quadratic form A_{σ_1}.

As $w \in C^1(\bar{B}(R) \setminus \bar{B}(1))$ for all $1 < R < 2$ and F_1 is C^1-smooth in $\bar{B}(2) \setminus \bar{B}(1))$, we can again use the chain rule, the area formula, and the monotone convergence theorem to obtain

$$\begin{aligned} L_{\sigma_1}[h] \le A_{\sigma_1}[v] &= \lim_{R\to 2}\lim_{\rho\to 0}\int_{B(R)\setminus\bar{B}(\rho)} \nabla v\cdot\sigma_1\nabla v\,dm(x) \\ &= \lim_{R\to 2}\lim_{\rho\to 0}\int_{F_1(B(R)\setminus\bar{B}(\rho))} \nabla w\cdot\gamma_1\nabla w\,dm(x) = L_{\gamma_1}[h]. \end{aligned} \tag{2-38}$$

Next, consider the inequality opposite to (2-38). We have by (2-35) and (2-36) that

$$L_{\sigma_1}[h] \ge \lim_{r\to 1} L_{\sigma_r}[h] = \lim_{r\to 1} L_{\gamma_r}[h] = L_{\gamma_1}[h]. \tag{2-39}$$

The above inequalities prove the claim. □

3. Complex geometric optics solutions

In what follows, we assume that $\mathcal{A}$ satisfies the almost linear growth condition (1-25).

3A. ***Existence and properties of the complex geometric optics solutions.*** Let us start with the observation that if $\sigma_0 \in \Sigma(\Omega_0)$ is a conductivity in a smooth simply connected domain $\Omega_0 \subset \mathbb{C}$, and σ_1 is a conductivity in a larger smooth domain Ω_1 which coincides with σ_0 in Ω_0 and is 1 in $\Omega_1 \setminus \Omega_0$, then L_{σ_0} determines L_{σ_1} by the formula

$$L_{\sigma_1}[h] = \inf\left\{ \int_{\Omega_1\setminus\Omega_0} |\nabla v|^2\,dm(z) + L_{\sigma_0}[v|_{\partial\Omega_0}] \;\middle|\; v \in W^{1,2}(\Omega_1 \setminus \overline{\Omega_0}),\; v|_{\partial\Omega_1} = h \right\}.$$

This observation implies that we may consider inverse problems by assuming that the conductivity σ is the identity near $\partial\Omega$ without loss of generality. Also, we may assume that $\Omega = \mathbb{D}$, which we do below. We note that boundary values of the isotropic conductivity can also be directly determined from Λ_σ; see [Alessandrini 1990].

The main result of this section is the following uniqueness and existence theorem for the complex geometrical optics solutions.

Theorem 3.1. *Let $\sigma \in \Sigma_{\mathcal{A}}(\mathbb{C})$ be a conductivity such that $\sigma(x) = 1$ for $x \in \mathbb{C} \setminus \Omega$. Then for every $k \in \mathbb{C}$, there is a unique solution $u(\,\cdot\,, k) \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$, where P is given in* (1-26), *for*

$$\nabla_z \cdot \sigma(z)\nabla_z u(z,k) = 0 \qquad \text{in } \mathbb{C}, \tag{3-1}$$

$$u(z,k) = e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{as } |z| \to \infty. \tag{3-2}$$

We point out that the regularity $u \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$ is optimal in the sense that the standard slightly stronger assumption $u \in W^{1,2}_{\mathrm{loc}}(\mathbb{C})$ would not be valid for the solutions; see [Astala et al. 2009, Section 20.4.6].

We prove Theorem 3.1 in several steps. Recalling the reduction to the Beltrami equation (2-7), we start with the following lemma, where we define

$$B_{\mathcal{A}}(\mathbb{D}) = \left\{ \mu \in L^\infty(\mathbb{C}) \;\middle|\; \operatorname{supp}(\mu) \subset \bar{\mathbb{D}},\; 0 \le \mu(x) < 1 \text{ a.e., and } \int_{\mathbb{D}} \exp(\mathcal{A}(K_\mu(z)))\,dm(z) < \infty \right\}.$$

Lemma 3.2. *Assume that* $\mu \in B_{\mathcal{A}}(\mathbb{D})$ *and* $f \in W^{1,P}_{\rm loc}(\mathbb{C})$ *satisfies*

$$\bar\partial f(z) = \mu(z)\,\partial f(z) \qquad \text{for a.e. } z \in \mathbb{C}, \tag{3-3}$$

$$f(z) = \beta e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z| \to \infty, \tag{3-4}$$

where $\beta \in \mathbb{C} \setminus \{0\}$ *and* $k \in \mathbb{C}$*. Then*

$$f(z) = \beta e^{ik\Phi(z)}, \tag{3-5}$$

where $\Phi \in W^{1,P}_{\rm loc}(\mathbb{C})$ *is a homeomorphism* $\Phi : \mathbb{C} \to \mathbb{C}$, $\bar\partial\Phi(z) = 0$ *for* $|z| > 1$, $K(z, \Phi) = K(z, f)$ *for a.e.* $z \in \mathbb{C}$, *and*

$$\Phi(z) = z + \mathcal{O}\Big(\frac{1}{z}\Big) \quad \text{for } |z| \to \infty. \tag{3-6}$$

Proof. By Theorem 2.2, we have for f the Stoilow factorization $f = h \circ \Phi$, where $h : \mathbb{C} \to \mathbb{C}$ is a holomorphic function and Φ is the principal solution of (3-3). This and the formulae (3-4) and (3-6) imply

$$\frac{h(\Phi(z))}{\beta e^{ik\Phi(z)}} = \frac{f(z)}{\beta e^{ik\Phi(z)}} \to 1 \quad \text{when } |z| \to \infty.$$

Thus, $h(\zeta) = \beta e^{ik\zeta}$ for all $\zeta \in \mathbb{C}$, and f has the representation (3-5). The claimed properties of Φ follow from the formula (3-5) and the similar properties of f. □

Next we consider case where $\beta = 1$. Below we will use the fact that if $\Phi : \mathbb{C} \to \mathbb{C}$ is a homeomorphism such that $\Phi \in W^{1,1}_{\rm loc}(\mathbb{C})$, we have $\Phi(z) - z = o(1)$ as $z \to \infty$ and that if Φ is analytic outside the disc $\bar B(r)$, $r > 0$, then by [Astala et al. 2009, Theorem 2.10.1 and (2.61)],

$$|\Phi(z)| \le |z| + 3r \quad \text{for } z \in \mathbb{C} \qquad \text{and} \qquad |\Phi(z) - z| \le r \quad \text{for } |z| > 2r. \tag{3-7}$$

In particular, the map Φ defined in Lemma 3.2 satisfies this with $r = 1$.

Lemma 3.3. *Assume that* $\nu, \mu : \mathbb{C} \to \mathbb{C}$ *are measurable functions satisfying*

$$\mu(z) = \nu(z) = 0 \quad \text{for } z \in \mathbb{C} \setminus \mathbb{D}, \tag{3-8}$$

$$|\mu(z)| + |\nu(z)| < 1 \quad \text{for a.e. } z \in \mathbb{D}, \tag{3-9}$$

and that $K_{\mu,\nu}(z)$ *defined in* (2-9) *satisfies*

$$\int_{\mathbb{D}} \exp(\mathcal{A}(K_{\mu,\nu}(z)))\,dm(z) < \infty. \tag{3-10}$$

Then for $k \in \mathbb{C}$, *the equation*

$$\partial_{\bar z} f = \mu\,\partial_z f + \nu\,\overline{\partial_z f}, \quad z \in \mathbb{C}, \tag{3-11}$$

has at most one solution $f \in W^{1,P}_{\rm loc}(\mathbb{C})$ *satisfying*

$$f(z) = e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z| \to \infty. \tag{3-12}$$

Proof. Observe that we can write (3-11) in the form

$$\partial_{\bar z} f = \tilde\mu\, \partial_z f, \quad z \in \mathbb{C}, \tag{3-13}$$

where the coefficient $\tilde\mu$ is given by (2-13). Since $|\tilde\mu(z)| \le |\mu(z)| + |\nu(z)|$, we see that $\tilde\mu \in B_{\mathcal{A}}(\mathbb{D})$.

Next, assume (3-13) has two solutions f_1 and f_2 having the asymptotics (3-12). Let $\varepsilon > 0$ and consider the function

$$f_\varepsilon(z) = f_1(z) - (1+\varepsilon) f_2(z).$$

Then, $f_\varepsilon \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$, the function f_ε satisfies (3-11), and

$$f_\varepsilon(z) = -\varepsilon e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z| \to \infty.$$

By Lemma 3.2 and (3-7), there is $\Phi_\varepsilon(z)$ such that

$$f_\varepsilon(z) = f_1(z) - (1+\varepsilon) f_2(z) = -\varepsilon e^{ik\Phi_\varepsilon(z)}$$

and $|\Phi_\varepsilon(z)| \le |z| + 3$. Then for any $z \in \mathbb{C}$, we have that

$$f_1(z) - f_2(z) = \lim_{\varepsilon \to 0} f_\varepsilon(z) = 0.$$

Thus $f_1 = f_2$. □

3B. ***Proof of Theorem 3.1.*** In the following, we use general facts for *weakly monotone* mappings, and to this end, we recall some basic facts. Let $\Omega \subset \mathbb{C}$ be open and $u \in W^{1,1}(\Omega)$ be real-valued. We say that u is weakly monotone if both of the functions $u(x)$ and $-u(x)$ satisfy the maximum principle in the following weak sense: for any $a \in \mathbb{R}$ and relatively compact open sets $\Omega' \subset \Omega$,

$$\max(u(z) - a, 0) \in W^{1,1}_0(\Omega') \text{ implies that } u(z) \le a \text{ for a.e. } z \in \Omega';$$

see [Iwaniec and Martin 2001, Section 7.3]. We remark that if $f \in W^{1,1}_{\mathrm{loc}}(\Omega_1)$ and $f : \Omega_1 \to \Omega_2$ is a homeomorphism, where $\Omega_1, \Omega_2 \subset \mathbb{C}$ are open, the real part of f is weakly monotone. By [Astala et al. 2009, Lemma 20.5.8], if $f \in W^{1,1}(\Omega)$ is the solution of the Beltrami equation $\bar\partial f = \mu\, \partial f$ with a Beltrami coefficient μ satisfying $|\mu(z)| < 1$ for a.e. $z \in \mathbb{C}$, then the real and the imaginary parts of f are weakly monotone functions. An important property of weakly monotone functions is that their modulus of continuity can be estimated in an explicit way. Let $M(t) = M_P(t)$ be the P-modulus, that is, the function determined by the condition: for $M = M(t)$, we have

$$\int_1^{1/t} P(sM) \frac{ds}{s^3} = P(1) \quad \text{for all } t \in [0, \infty);$$

see (1-30) and [Iwaniec and Martin 2001, Section 7.5]. The function $M_P : [0, \infty) \to [0, \infty)$ is continuous at zero and $M_P(0) = 0$. Then by [Iwaniec and Martin 2001, Theorem 7.5.1], it holds that if $z', z \in \Omega$ satisfy $B(z, r) \subset \Omega$, $r < 1$, and $|z' - z| < r/2$, and $f \in W^{1,P}(\Omega)$ is a weakly monotone function, then for almost every $z, z' \in B(z, r)$, we have

$$|f(z') - f(z)| \le 32\pi r \, \|Df\|_{(P,r)} M_P\Big(\frac{|z-z'|}{2r}\Big), \tag{3-14}$$

where

$$\|\nabla f\|_{(P,r)} = \inf\Big\{\frac{1}{\lambda} \,\Big|\, \lambda > 0,\ \frac{1}{\pi r^2}\int_{B(z,r)} P(\lambda |Df(x)|)\, dm(z) \le P(1)\Big\}.$$

As we will see, this can be used to estimate the modulus of continuity of principal solutions of Beltrami equations corresponding to $\mu \in B_{\mathcal{A}}(\mathbb{D})$.

Below, we use the unimodular function e_k given by

$$e_k(z) = e^{i(kz+\bar k \bar z)}. \tag{3-15}$$

The following result shows the existence of the complex geometric solutions for degenerated conductivities.

Lemma 3.4. *Assume that μ and ν satisfy* (3-8)–(3-10) *and let $k \in \mathbb{C}\setminus\{0\}$. Then* (3-11) *has a solution $f \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$ satisfying the asymptotics* (3-12). *Moreover, this solution can be written in the form*

$$f(z) = e^{ik\varphi(z)}, \tag{3-16}$$

where $\varphi : \mathbb{C} \to \mathbb{C}$ is a homeomorphism satisfying the asymptotics $\varphi(z) = z + \mathcal{O}(z^{-1})$. Moreover, for $R > 1$,

$$\int_{B(R)} P(|D\varphi(x)|)\, dm(x) \le C_{\mathcal{A}}(R) \int_{B(R)} \exp(\mathcal{A}(K_{\mu,\nu}(z)))\, dm(z), \tag{3-17}$$

where $C_{\mathcal{A}}(R)$ depends on R and the weight function $\mathcal{A}$. In addition,

$$\bar\partial\varphi(z) = \mu(z)\, \partial\varphi(z) - \frac{\bar k}{k}\nu(z) e_{-k}(\varphi(z))\, \overline{\partial\varphi(z)} \quad \textit{for a.e. } z \in \mathbb{C}. \tag{3-18}$$

Proof. Let us approximate the functions μ and ν with functions

$$\mu_n(z) = \begin{cases} \mu(z) & \text{if } |\mu(z)| + |\nu(z)| \le 1 - \frac{1}{n}, \\ \frac{\mu(z)}{|\mu(z)|}(1-\frac{1}{n}) & \text{if } |\mu(z)| + |\nu(z)| > 1 - \frac{1}{n}, \end{cases} \tag{3-19}$$

$$\nu_n(z) = \begin{cases} \nu(z) & \text{if } |\mu(z)| + |\nu(z)| \le 1 - \frac{1}{n}, \\ \frac{\nu(z)}{|\nu(z)|}(1-\frac{1}{n}) & \text{if } |\mu(z)| + |\nu(z)| > 1 - \frac{1}{n}, \end{cases} \tag{3-20}$$

where $n \in \mathbb{Z}_+$. Consider the equations

$$\bar\partial f_n(z) = \mu_n(z)\partial f_n(z) + \nu_n(z)\overline{\partial f_n(z)} \quad \text{for a.e. } z \in \mathbb{C}, \tag{3-21}$$

$$f_n(z) = e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z| \to \infty. \tag{3-22}$$

By Lemma 3.3, equations (3-21)–(3-22) have at most one solution $f_n \in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$. The existence of the solutions can be seen as in the proof of [Astala et al. 2005, Lemma 3.5]; by [Astala et al. 2005, Lemma 3.2], solutions f_n for (3-21)–(3-22) can be constructed via the formula $f_n = h \circ g$, where g is the principal solution of $\bar\partial g = \hat\mu \partial g$, constructed in Theorem 2.2, and h is the solution of

$$\bar\partial h = (\hat\nu \circ g^{-1})\, \overline{\partial h}, \quad h(z) = e^{ikz}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big)$$

constructed in [Astala and Päivärinta 2006, Theorem 4.2], where $\hat\nu = (1+\nu_n)\bar\mu_n$ and $\hat\mu = \mu_n(1+\hat\nu)$, and moreover, it holds that $f_n \in W^{1,2}_{\mathrm{loc}}(\mathbb{C})$.

Let us now define the coefficient $\tilde\mu$ according to formula (2-13), and define an approximative coefficient $\tilde\mu_n$ using formula (2-13), where μ and ν are replaced by μ_n and ν_n and f by f_n. We can write (3-21) in the form

$$\bar\partial f_n(z) = \tilde\mu_n(z)\, \partial f_n(z) \quad \text{for a.e. } z \in \mathbb{C}, \tag{3-23}$$

where $|\tilde\mu_n| \le 1 - n^{-1}$.

By (3-22), (3-23), and Lemma 3.2, the function f_n can be written in the form

$$f_n(z) = e^{ik\varphi_n(z)}, \tag{3-24}$$

where φ_n is a homeomorphism, $\bar\partial\varphi_n(z) = 0$ for $|z| > 1$, $K(z, \varphi_n) = K(z, f_n)$ for a.e. $z \in \mathbb{C}$, and

$$\varphi_n(z) = z + \mathcal{O}\Big(\frac{1}{z}\Big) \quad \text{for } |z| \to \infty. \tag{3-25}$$

Then

$$|\bar\partial f_n(z)| = |\tilde\mu_n(z)|\,|\partial f_n(z)| \le |\tilde\mu(z)|\,|\partial f_n(z)|.$$

Let us consider next $a, b > 0$ and $0 \le t \le (ab)^{1/2}$. Using the definition (1-26) of $P(t)$, we see that

$$\begin{aligned} P(t) &\le \exp(\mathcal{A}(a)) && \text{for } t^2 \le e^{\mathcal{A}(a)},\\ P(t) &\le \frac{ab}{\mathcal{A}^{-1}(\log\exp(\mathcal{A}(a)))} = b && \text{for } t^2 > e^{\mathcal{A}(a)}, \end{aligned}$$

which imply the inequality $P(t) \le b + \exp(\mathcal{A}(a))$. Due to the distortion equality (2-15), we can use this for $a = K(z, \varphi_n)$, $b = J(z, \varphi_n)$, and $t = |D\varphi_n(z)|$ and obtain

$$P(|D\varphi_n(z)|) \le J(z, \varphi_n) + \exp(\mathcal{A}(K(z, \varphi_n))). \tag{3-26}$$

Then, we see using (3-7) and the fact that φ_n is a homeomorphism that

$$\begin{aligned} \int_{B(R)} P(|D\varphi_n(z)|)\, dm(z) &\le \int_{B(R)} J(z, \varphi_n)\, dm(z) + \int_{B(R)} e^{\mathcal{A}(K(z,\varphi_n))}\, dm(z)\\ &\le m(\varphi_n(B(R))) + \int_{B(R)} \exp(\mathcal{A}(K_{\tilde\mu}(z)))\, dm(z)\\ &\le \pi(R+3)^2 + \int_{B(R)} \exp(\mathcal{A}(K_{\tilde\mu}(z)))\, dm(z) \end{aligned} \tag{3-27}$$

is finite by the assumption (3-10). We emphasize that the fact that φ_n is a homeomorphism is the essential fact which together with the inequality (3-26) yields the Orlicz estimate (3-27).

The estimate (3-27) together with the inequality (3-14) implies that the functions φ_n have uniformly bounded modulus of continuity in all compact sets of $\mathbb{C}$. Moreover, by (3-7), $|\varphi_n(z)| \le |z| + 3$.

Next we consider the Beltrami equation for φ. To this end, let $\psi \in C_0^\infty(\mathbb{C})$ and $R > 1$ be so large that $\mathrm{supp}(\psi) \subset B(R)$. Since the family $\{\varphi_n\}_{n=1}^\infty$ is uniformly bounded in the space $W^{1,P}(B(R))$ and $W^{1,P}(B(R)) \subset W^{1,q}(B(R))$ for some $q > 1$, we see that there is a subsequence φ_{n_j} that converges weakly in $W^{1,q}(B(R))$ to some limit φ when $j \to \infty$. Let us denote

$$\kappa_n(z) = -\frac{\bar{k}}{k}\nu_n(z)e_{-k}(\varphi_n(z)), \quad \kappa(z) = -\frac{\bar{k}}{k}\nu(z)e_{-k}(\varphi(z)).$$

Moreover, functions φ_n are uniformly bounded and have a uniformly bounded modulus of continuity in compact sets by (3-14) and thus by the Arzelà–Ascoli theorem, there is a subsequence, denoted also by φ_{n_j}, that converges uniformly to some function φ' in $B(R)$ for all $R > 1$. As φ_{n_j} converges in $C(\bar{B}(R))$ uniformly to φ' and weakly in $W^{1,q}(B(R))$ to φ, we see using convergence in distributions that $\varphi' = \varphi$. Thus, we see that

$$\lim_{j\to\infty} e_{-k}(\varphi_{n_j}(z)) = e_{-k}(\varphi(z)) \quad \text{uniformly for } z \in B(R),$$

and by the dominated convergence theorem $\kappa_n \to \kappa$ in $L^p(B(R))$, where $1/p + 1/q = 1$.

As $\varphi_n : \mathbb{C} \to \mathbb{C}$ is a homeomorphism and $\varphi_n \in W^{1,1}_{\mathrm{loc}}(\mathbb{C})$, we can use chain rules (2-10) a.e. by the Gehring–Lehto theorem (see [Astala et al. 2009, Corollary 3.3.3]) and see using (3-21) and (3-24) that

$$\bar{\partial}\varphi_n(z) = \mu_n(z)\,\partial\varphi_n(z) - \frac{\bar{k}}{k}\nu_n(z)\,e_{-k}(\varphi_n(z))\,\overline{\partial\varphi_n(z)} \quad \text{for a.e. } z \in \mathbb{C}. \tag{3-28}$$

Recall that there is a convex function $\Phi : [0,\infty) \to [0,\infty)$ such that $\Phi(t) \le P(t) + c_0 \le 2\Phi(t)$. By [Attouch et al. 2006, Theorem 13.1.2], the map

$$\phi \mapsto \int_{B(R)} \Phi(|D\phi(x)|)\,dm(x)$$

is weakly lower semicontinuous in $W^{1,1}(B(R))$. By (3-27), the integral of $\Phi(|D\varphi_n|)$ is uniformly bounded in $n \in \mathbb{Z}_+$ over any disc $B(R)$. In particular, this yields that $\varphi \in W^{1,P}(B(R))$ for $R > 1$ and that (3-17) holds.

Furthermore, as $|\varphi(z)| \le |z| + 3$, this yields that

$$f(z) := e^{ik\varphi(z)} \in W^{1,P}_{\mathrm{loc}}(\mathbb{C}). \tag{3-29}$$

Next define $\varphi_n(\infty) = \varphi(\infty) = \infty$. As φ_n and φ are conformal at infinity, we see using the Cauchy formula for $(\varphi_n(1/z) - \varphi(0))^{-1}$ that

$$\varphi(z) = z + \mathcal{O}\Big(\frac{1}{z}\Big) \quad \text{for } |z| \to \infty. \tag{3-30}$$

As $D\varphi_{n_j}$ converges weakly in $L^q(B(R))$ to $D\varphi$ and their norms are uniformly bounded, we have

$$\begin{aligned}\Big|\int_{\mathbb{C}}(\bar{\partial}\varphi - \mu\,\partial\varphi - \kappa\,\overline{\partial\varphi})\psi\,dm(z)\Big| &= \lim_{j\to\infty}\Big|\int_{\mathbb{C}}(\bar{\partial}\varphi_{n_j} - \mu\,\partial\varphi_{n_j} - \kappa\,\overline{\partial\varphi_{n_j}})\psi\,dm(z)\Big| \\ &\le \lim_{j\to\infty}\Big|\int_{\mathbb{C}} i\big((\mu_{n_j} - \mu)\,\partial\varphi_{n_j} + (\kappa_{n_j} - \kappa)\,\overline{\partial\varphi_{n_j}}\big)\psi\,dm(z)\Big| \\ &\le \lim_{j\to\infty}\big(\|\mu_{n_j} - \mu\|_{L^p(B(1))} + \|\kappa_{n_j} - \kappa\|_{L^p(B(1))}\big)\|\partial\varphi_{n_j}\|_{L^q(B(1))}\|\psi\|_{L^\infty(B(1))} = 0.\end{aligned}$$

This implies that $\varphi(z)$ satisfies (3-18).

Next we show that φ is a homeomorphism As $K(z)=K_{\nu,\mu}\in L^1_{\mathrm{loc}}(\mathbb{C})$, we have $K(z;\varphi_n)\in L^1_{\mathrm{loc}}(\mathbb{C})$; thus by (2-16), the inverse maps φ_n^{-1} satisfy $\varphi_n^{-1}\in W^{1,2}_{\mathrm{loc}}(\mathbb{C};\mathbb{C})$ and for all $R>1$, the norms $\|\varphi_n^{-1}\|_{W^{1,2}(B(R))}$, $n\in\mathbb{Z}_+$, are uniformly bounded. Thus by the formula (3-14), the family $(\varphi_n^{-1})_{n=1}^{\infty}$ has a uniform modulus of continuity in compact sets. Hence, we see that there is a continuous function $\psi:\mathbb{C}\to\mathbb{C}$ such that $\varphi_n^{-1}\to\psi$ uniformly on compact sets when $n\to\infty$. As φ_n are conformal at infinity, we see, using again the Cauchy formula, that $\varphi_{n_j}^{-1}\to\psi$ uniformly on the Riemann sphere $\mathbb{S}^2$ as $j\to\infty$. Then,

$$\psi\circ\varphi(z)=\lim_{j\to\infty}\varphi_{n_j}^{-1}(\varphi(z))=\lim_{j\to\infty}\varphi_{n_j}^{-1}(\varphi_{n_j}(z))=z,$$

which implies that $\varphi:\mathbb{S}^2\to\mathbb{S}^2$ is a continuous injective map and hence a homeomorphism.

As $\varphi:\mathbb{C}\to\mathbb{C}$ is a homeomorphism and $\varphi\in W^{1,1}_{\mathrm{loc}}(\mathbb{C})$, we can, by the Gehring–Lehto theorem, use chain rules (2-10) a.e. and see using (3-18) that $f(z)=e^{ik\varphi(z)}$ satisfies (3-11). By (3-30), $f(z)$ satisfies the asymptotics (3-12). This proves the claim. □

The above uniqueness and existence results have now proven Theorem 3.1.

4. Inverse conductivity problem with degenerate isotropic conductivity

In this section, we consider exponentially integrable scalar conductivities σ. In particular, we assume that σ is 1 in an open set containing $\mathbb{C}\setminus\mathbb{D}$ and its ellipticity function $K(z)=K_\sigma(z)$ of the conductivity σ satisfies an Orlicz space estimate

$$\int_{B(R_1)}\exp\big(\exp(qK(x))\big)\,dm(x)\le C_0\quad\text{for some } C_0,q>0,\tag{4-1}$$

with $R_1=1$. Note that by the John–Nirenberg lemma, (4-1) is satisfied if

$$\exp(qK(x))\in\mathrm{BMO}(\mathbb{D})\quad\text{for some } q>0.\tag{4-2}$$

As noted before, we may assume without loss of generality that Ω is the unit disc $\mathbb{D}$.

4A. ***Estimates for principal solutions in Orlicz spaces.*** Let us consider next the principal solution of the Beltrami equation

$$\bar\partial\Phi(z)=\mu(z)\,\partial\Phi(z),\quad z\in\mathbb{C},\tag{4-3}$$

$$\Phi(z)=z+O\Big(\frac{1}{z}\Big)\qquad\text{when } |z|\to\infty.\tag{4-4}$$

To this end, let $R_0\ge1$,

$$B^p_{\exp,N}(B(R_0))=\left\{\mu:\mathbb{C}\to\mathbb{C}\;\middle|\;|\mu(z)|<1\text{ for a.e. } z,\ \mathrm{supp}(\mu)\subset B(R_0)\text{ and }\int_{B(R_0)}\exp(pK_\mu(z))\,dm(z)\le N\right\}$$

and

$$B^p_{\exp}(B(R_0))=\bigcup_{N>0}B^p_{\exp,N}(B(R_0)).$$

The reason that we use the radius R_0 is to be able to apply the obtained results for the inverse function of the solution of the Beltrami equation satisfying another Beltrami equation with modified coefficients; see (4-45).

Assume that $p > 2$ and $\mu \in B^p_{\exp}(B(R_0))$. Then by [Astala et al. 2010, Theorem 1.1], we have the L^2-estimate

$$\|(\mu S)^m \mu\|_{L^2(\mathbb{C})} \le C(p,\beta) m^{-\beta/2} \int_{B(R_0)} \exp(pK_\mu(z))\, dm(z), \quad 2 < \beta < p, \tag{4-5}$$

where

$$S\phi(z) = \frac{1}{\pi} \int_{\mathbb{C}} \frac{\partial_w \phi(w)}{w - z}\, dm(w),$$

is the Beurling operator. Below, we use the operator $S_{B(R_0)}\phi = S\phi|_{B(R_0)}$. In particular, as Φ satisfies

$$\bar\partial\Phi = \bar\partial(\Phi - z) = \mu\, \partial(\Phi - z) + \mu = \mu S\, \bar\partial(\,\Phi - z) + \mu = \mu S\, \bar\partial\Phi + \mu,$$

(4-5) yields

$$\bar\partial\Phi = \sum_{m=0}^{\infty} (\mu S)^m \mu, \tag{4-6}$$

where the series converges in $L^2(\mathbb{C})$. To analyze the convergence more precisely, we need a refinement of the L^p-scale. In particular, we will use the Orlicz spaces $X^{j,q}(S)$, $j \in \mathbb{Z}_+$, $q \in \mathbb{R}$, $S \subset \mathbb{C}$ that are defined by

$$u \in X^{j,q}(S) \quad \text{if and only if} \quad \int_S M_{j,q}(u(x))\, dm(x) < \infty, \tag{4-7}$$

where

$$M_{j,q}(t) = |t|^j \log^q(e + |t|). \tag{4-8}$$

We use shorthand notations $X^q(S) = X^{2,q}(S)$ and $M_q(t) = M_{2,q}(t)$. Although (4-7)–(4-8) do not define a norm in $X^{j,q}(S)$, there is an equivalent norm

$$\|u\|_{X^{j,q}(S)} = \sup_v \left\{ \int_S |u(x)v(x)|\, dm(x) \;\middle|\; \int_D G_{j,q}(|v(x)|)\, dm(x) \le 1 \right\}, \tag{4-9}$$

where $G_{j,q}(t)$ is such a function that $(M_{j,q}, G_{j,q})$ are a Young complementary pair (see the Appendix) and, in particular, the following lemma holds.

Lemma 4.1. *Let $j = 1, 2, \ldots,$ and $q \ge 0$.*

(i) *We have*

$$\int_{B(R_0)} M_{j,q}(u(x))\, dm(x) \le 2\|u\|^j_{X_{j,q}(B(R_0))} \log^q(e + \|u\|_{X^{j,q}(B(R_0))}).$$

(ii) *For*

$$\phi(t) = t^{1/j}(1 + 2\log^q(e + t^{-1/j})),$$

we have

$$\|u\|_{X^{j,q}(B(R_0))} \le \phi\left(\int_{B(R_0)} M_{j,q}(u(x))\, dm(x) \right).$$

Proof. (i) Let us denote $M(t) = M_{j,q}(t)$. For this function, we use the equivalent norms $\|u\|_M$ and $\|u\|_{(M)}$ defined in the Appendix. To show the claim, we use the inequality

$$\log(e + st) \le 2\log(e + s)\log(e + t), \quad t, s \ge 0. \tag{4-10}$$

Let us consider the function $w \in X^{j,q}(B(R_0))$. By (4-10) we have, for $k > 0$, that

$$\int_{B(R_0)} M_{j,q}(kw)\,dm = k^j \int_{B(R_0)} |w|^j \log^q(e + k|w|)\,dm \le 2k^j \log^q(e+k) \int_{B(R_0)} M_{j,q}(w)\,dm. \tag{4-11}$$

A function $u \in X^{j,q}(B(R_0))$ can be written as $u = kw$, where $k = \|u\|_{(M)}$ and $\|w\|_{(M)} = 1$. Then by (A-5)–(A-6), we have

$$\int_{B(R_0)} M_{j,q}(w)\,dm = 1,$$

and hence (4-11) and (A-4) yield the claim (i).

(ii) Using (4-11) and the definition (A-2) of the Orlicz norm, we see that for all $k > 0$,

$$\|u\|_{X^{j,q}(B(R_0))} \le \frac{1}{k}\biggl(1 + \int_{B(R_0)} M_{j,q}(ku)\,dm\biggr) \le \frac{1}{k}\biggl(1 + 2k^j \log^q(e+k) \int_{B(R_0)} M_{j,q}(u)\,dm\biggr).$$

Let $T = \int_{B(R_0)} M_{j,q}(u)\,dm$. Substituting $k = T^{-1/j}$ above, we obtain (ii). □

Theorem 4.2. *Assume that $\mu \in B^p_{\exp}(B(R_0))$, $2 < p < \infty$. Then the equations* (4-3)–(4-4) *have a unique solution $\Phi \in W^{1,1}_{\mathrm{loc}}(\mathbb{C})$ which, for $0 \le q \le p/4$, satisfies*

$$\bar\partial \Phi \in X^q(\mathbb{C}) \tag{4-12}$$

and the series (4-6) *converges in $X^q(\mathbb{C})$. The convergence of the series* (4-6) *in $X^q(\mathbb{C})$ is uniform for $\mu \in B^p_{\exp,N}(B(R_0))$ with any $N > 0$. Moreover, for $\mu \in B^p_{\exp,N}(B(R_0))$, the Jacobian $J_\Phi(z)$ of Φ satisfies*

$$\|J_\Phi\|_{X^{1,q}(B(R_0))} \le C, \tag{4-13}$$

where C depends only on p, q, N, and R_0. Moreover, let $s > 2$ and assume that $\mu_m, \tilde\mu_m \in B^p_{\exp,N}(B(R_0))$ and $0 \le q \le p/4$. Then we have the following implication:

$$\lim_{m\to\infty} \|\mu_m - \tilde\mu_m\|_{L^s(B(R_0))} = 0 \quad \Rightarrow \quad \lim_{m\to\infty} \|\bar\partial \Phi_m - \bar\partial \tilde\Phi_m\|_{X^q(\mathbb{C})} = 0, \tag{4-14}$$

where Φ_m and $\tilde\Phi_m$ are the solutions of (4-3)–(4-4) *corresponding to μ_m, $\tilde\mu_m$, respectively.*

Proof. Let $\Phi^\lambda(z)$, where $|\lambda| \le 1$, $z \in \mathbb{C}$, be the principal solution corresponding to the Beltrami coefficient $\lambda\mu$, that is, the solution with the Beltrami equation (2-17)–(2-18) with coefficient $\lambda\mu$. These solutions, in particular $\Phi^\lambda = \Phi^1$, exist and are unique by Theorem 2.2. It follows from [Astala et al. 2010, Theorems 1.1 and 5.1] that the Jacobian determinant $J\Phi^\lambda(z)$ of Φ^λ satisfies

$$\int_{B(R_0)} J\Phi^\lambda \log^{2q}(e + J\Phi^\lambda)\,dm(z) \le C < \infty, \tag{4-15}$$

where C is independent of λ and $\mu \in B^p_{\exp,N}(B(R_0))$ and depends only on N, p, and q. Thus (4-13) follows from Lemma 4.1(ii).

We showed already that when $p > 2$, we have $\bar\partial\Phi \in L^2(\mathbb{C})$ and that the series (4-6) converges in $L^2(\mathbb{C})$. To show the convergence of (4-6) in $X^q(\mathbb{C})$ and to prove (4-14), we present a few lemmas in terms of Orlicz spaces $X^q(B(R_0))$ and the function M_q defined in (4-8). Note that as μ vanishes in $\mathbb{C} \setminus B(R_0)$,

$$\|(\mu S)^n\mu\|_{X^q(\mathbb{C})} = \|(\mu S)^n\mu\|_{X^q(B(R_0))}.$$

Lemma 4.3. *Let $N \in \mathbb{Z}_+$, $2 < 2q < \beta < p$, and $\mu \in B^p_{\exp,N}(B(R_0))$. Then*

$$\int_{B(R_0)} M_q(\psi_n(x))\,dm(x) \le cn^{-(\beta-q)} < cn^{-q}, \tag{4-16}$$

where $\psi_n = (\mu S)^n\mu$ and $c > 0$ depends only on N, p, β, and q.

Proof. Let $E_n = \{z \in B(R_0) : |\psi_n(z)| \ge A^n\}$, where $A > 1$ is a constant to be chosen later. By (4-5),

$$\|\psi_n\|_{L^2(B(R_0))} \le C_{N,\beta,p}n^{-\beta/2}. \tag{4-17}$$

Thus

$$|E_n| \le C^2_{N,\beta,p}A^{-2n}n^{-\beta}. \tag{4-18}$$

Using (4-17), we obtain

$$\int_{B(R_0)\setminus E_n} |\psi_n|^2 \log^q(e+|\psi_n|)\,dm \le \|\psi_n\|^2_{L^2(B(R_0))} \log^q(e+A^n) \le C_1n^{-\beta+q}, \tag{4-19}$$

where $C_1 = C^2_{N,\beta,p}\log^q(e+A)$.

The principal solution corresponding to the Beltrami coefficient $\lambda\mu$ can be written in the form

$$\Phi^\lambda(z) = z + \frac{1}{\pi}\int_{\mathbb{C}} \frac{\bar\partial\Phi^\lambda(w)}{w-z}\,dm(w), \qquad \bar\partial\Phi^\lambda = (I-\lambda\mu S)^{-1}(\lambda\mu).$$

Expanding $\bar\partial_z\Phi^\lambda(z)$ as a power series in λ, we see that by (4-6) we can write, using any $0 < \rho < 1$,

$$\chi_{E_n}(z)\psi_n(z) = \frac{1}{2\pi i}\int_{|\lambda|=\rho} \lambda^{-n-2}\chi_{E_n}(z)\,\bar\partial_z\Phi^\lambda(z)\,d\lambda.$$

This gives

$$\|\chi_{E_n}\psi_n\|_{X^q(B(R_0))} \le \rho^{-(n+2)} \sup_{|\lambda|=\rho} \|\chi_{E_n}\bar\partial_z\Phi^\lambda\|_{X^q(B(R_0))}. \tag{4-20}$$

Using the facts that $|\lambda| = \rho$ and that the Beltrami coefficient Φ^λ is bounded by $|\lambda|$, we have, by the distortion equality (2-15), that

$$|\bar\partial_z\Phi^\lambda(z)|^2 \le \rho^2(1-\rho^2)^{-1}J\Phi^\lambda(z).$$

Hence,

$$I := \int_{E_n} M_q(\bar\partial_z\Phi^\lambda(z))\,dm(z) \le \frac{\rho^2}{1-\rho^2}\int_{E_n} J\Phi^\lambda \log^q\left(e + \left(\frac{\rho^2}{1-\rho^2}J\Phi^\lambda\right)^{1/2}\right)dm.$$

Next, let $\widehat{C}$ denote a generic constant which is a function of N, β, p, q and ρ but not of A. The above implies by (4-10), (4-15), and the inequality $\log(e+t^{1/2}) \le 1+\log(e+t)$, $t \ge 0$, that

$$\begin{aligned} I &\le \widehat{C}\int_{E_n} J\Phi^\lambda(z)(1+\log(e+J\Phi^\lambda))^q\,dm(z) \\ &\le \widehat{C}\Big(\int_{E_n} J\Phi^\lambda(z)\,dm(z)\Big)^{1/2}\Big(\int_{E_n} J\Phi^\lambda(z)(1+\log(e+J\Phi^\lambda(z)))^{2q}\,dm(z)\Big)^{1/2} \\ &\le \widehat{C}\Big(\int_{E_n} J\Phi^\lambda(z)\,dm(z)\Big)^{1/2}. \end{aligned}$$

By the area distortion theorem from [Astala 1994], as formulated in [Astala et al. 2009, Theorem 13.1.4], we have

$$\int_{E_n} J\Phi^\lambda(z)\,dm(z) \le |\Phi^\lambda(E_n)| \le \widehat{C}|E_n|^{1/M} \le \widehat{C}A^{-2n/M},$$

where $M=(1+\rho)/(1-\rho)>1$, and thus $I \le \widehat{C}A^{-n/M}$. By Lemma 4.1(ii), we also have the estimate

$$\|\chi_{E_n}\bar\partial\Phi^\lambda\|_{X^q} \le \widehat{C}A^{-n/M}.$$

Taking $\rho > e^{-1/2}$ and $A=e^M$, we see using (4-20) and Lemma 4.1 again that

$$\int_{E_n} M_q(\psi_n)\,dm(z) \le \widehat{C}e^{-n/2}$$

for sufficiently large $n \in \mathbb{Z}_+$. Thus the assertion follows from (4-19). □

Lemmas 4.1(ii) and 4.3 and the fact that μ vanishes outside $B(R_0)$ yield that for $q>1$ and $p>2q$, there is an $N>0$ such that the series (4-6) converges in $X^q(\mathbb{C})$, and moreover, convergence of the series (4-6) is uniform for $\mu \in B^p_{\exp,N}(B(R_0))$. Thus to prove Theorem 4.2 it remains to show (4-14).

Lemma 4.4. *Let* $2<2q<p$, $N>0$, $2<\beta<p$, $s>2$, $\mu,\nu \in B^p_{\exp,N}(B(R_0))$, *and* $B_n=(\mu S)^n\mu-(\nu S)^n\nu$. *Then*

$$\sup_{n\in\mathbb{Z}_+}\int_{\mathbb{C}} M_q(B_n(x))\,dm(x) \le C, \tag{4-21}$$

where $C>0$ *depends only on* N, p, *and* q. *Moreover, there is* $T>1$ *such that*

$$\|B_n\|_{L^2(\mathbb{C})} \le C_{N,\beta,p,s,T}\min\big(nT^n\|\mu-\nu\|_{L^s(B(R_0))},\, n^{-\beta/2}\big). \tag{4-22}$$

Proof. Lemmas 4.1 and 4.3 yield (4-21). Next, let us observe that for $z\in\mathbb{C}$,

$$B_n(z)=(\mu S)^n\mu-(\nu S)^n\nu=\sum_{j=0}^n A_j(z), \quad A_j(z)=(\mu S)^j(\mu-\nu)(S\nu)^{n-j}\chi_{B(R_0)}.$$

As $\|\nu\|_{L^\infty}\le 1$ and $\|S\|_q := \|S\|_{L^q(\mathbb{C})\to L^q(\mathbb{C})} < \infty$ for $1<q<\infty$, we have that

$$\begin{aligned} \int_{\mathbb{C}} |A_j(z)|^q\,dm(z) &\le (\|S\|_q^q)^j \int_{B(R_0)} |\mu(z)-\nu(z)|^q\,|((S\nu)^{n-j}\chi_{B(R_0)})(z)|^q\,dm(z) \\ &\le \|S\|_q^{jq}\Big(\int_{B(R_0)}|\mu(z)-\nu(z)|^{q\rho}dm(z)\Big)^{1/\rho}\Big(\int_{B(R_0)}|((S\nu)^{n-j}\chi_{B(R_0)})(z)|^{q\rho'}dm(z)\Big)^{1/\rho'}, \end{aligned}$$

where $\rho^{-1} + (\rho')^{-1} = 1$ and $1 < \rho < \infty$. Thus

$$\|A_j(z)\|_{L^q(\mathbb{C})} \le (\|S\|_q)^j \|\mu - \nu\|_{L^{\rho q}(B(R_0))} (\|S\|_{q\rho'}^q)^{n-j} \|\nu\|_{L^{q\rho'}(B(R_0))}^q,$$

where $\|\nu\|_{L^{q\rho'}(B(R_0))} \le \pi R_0^2$. Thus by choosing $q = 2$ and ρ so that $s = q\rho > 2$ yielding $q\rho' = 2s/(s-2)$, we obtain

$$\|(\mu S)^n \mu - (\nu S)^n \nu\|_{L^2(\mathbb{C})} \le (n+1)\pi^2 R_0^4 (1 + \|S\|_{(2s/(s-2))}^2)^n \|\mu - \nu\|_{L^s(B(R_0))}.$$

This and (4-5) show that (4-22) is valid. □

Now we are ready to prove (4-14), which finishes the proof of Theorem 4.2. Let

$$B_{n,m} = (\mu_m S)^n \mu_m - (\tilde{\mu}_m S)^n \tilde{\mu}_m.$$

By the Schwarz inequality, we have that (4-21), (4-22) and Lemma 4.1 yield

$$\begin{aligned}\int_{B(R_0)} M_q(B_{n,m}(z))\, dm(z) &\le \int_{B(R_0)} |B_{n,m}|^2 \log^q(e + |B_{n,m}|)\, dm(z) \\ &\le \left(\int_{B(R_0)} M_{2q}(B_{n,m}(z))\, dm(z)\right)^{1/2} \|B_{n,m}\|_{L^2(B(R_0))} \\ &\le C \min\big(nT^n \|\mu_m - \tilde{\mu}_m\|_{L^s(B(R_0))}, n^{-\beta/2}\big), \qquad (4\text{-}23)\end{aligned}$$

where C depends only on q, p, β, s, T, and N.

Let $\varepsilon > 0$. As μ_m and $\tilde{\mu}_m$ vanish outside $B(R_0)$,

$$\|\bar{\partial}\Phi_m - \bar{\partial}\tilde{\Phi}_m\|_{X^q(\mathbb{C})} = \|\bar{\partial}\Phi_m - \bar{\partial}\tilde{\Phi}_m\|_{X^q(B(R_0))} \le \sum_{n=0}^{\infty} \|B_{n,m}\|_{X^q(B(R_0))}.$$

Thus by (4-23) and Lemma 4.1(ii), we can take $n_0 \in \mathbb{N}$ so large that for all m,

$$\sum_{n=n_0}^{\infty} \|B_{n,m}\|_{X^q(B(R_0))} \le \frac{\varepsilon}{2}.$$

Applying again (4-23) and Lemma 4.1(ii), we can choose $\delta > 0$ so that

$$\sum_{n=0}^{n_0-1} \|B_{n,m}\|_{X^q(B(R_0))} \le \frac{\varepsilon}{2} \quad \text{when } \|\mu_m - \tilde{\mu}_m\|_{L^s(B(R_0))} \le \delta.$$

This proves Theorem 4.2. □

Lemma 4.5. *Assume that K_μ corresponding to μ supported in $\mathbb{D}$ satisfies* (4-1) *with q, $C_0 > 0$ and $R_1 = 1$. Let Φ be the principal solution of the Beltrami equation corresponding to μ. Then for all β, $R > 0$, the inverse function $\Psi = \Phi^{-1} : \mathbb{C} \to \mathbb{C}$ of Φ satisfies*

$$\int_{B(R)} \exp(\beta K_\mu(\Psi(z)))\, dm(z) < C,$$

where C depends only on q, C_0, β, and R.

Proof. Since Φ satisfies the condition $\mathcal{N}$ by [Astala et al. 2010, Corollary 4.3], we may change variable in integration to see that

$$\int_{B(R)} \exp(\beta K_\mu(\Psi(z)))\, dm(z) = \int_{\Psi(B(R))} \exp(\beta K_\mu(w)) J_\Phi(w)\, dm(w). \tag{4-24}$$

Using (3-7) for the function Φ and $R > 3$, we see that $\Psi(B(R)) \subset \tilde{B} = B(\tilde{R})$, $\tilde{R} = R + 1$. By (4-1), $\exp(K_\mu(z)) \in L^q(\tilde{B})$ for all $q > 1$ and thus by (4-13), $J_\Phi \in X^{1,q}(B(R))$ for $R > 0$.

Let us next use properties of Orlicz spaces and the notations discussed in the Appendix using a Young complementary pair (F, G), where

$$F(t) = \exp(t^{1/p}) - 1$$

and $G(t)$ satisfies $G(t) = C_p t(\log(1 + C_p t))^p$ for $t > T_p$ with suitable $C_p, T_p > 0$; see [Krasnosel'skiĭ and Rutickiĭ 1961, Theorem I.6.1].

By using $u(z) = \exp(\beta K_\mu(z))$ and $v = J_\Phi(z)$, we obtain from Young's inequality (A-7) the inequality

$$\begin{aligned}\int_{B(R)} &\exp(\beta K_\mu(\Psi(z)))\, dm(z)\\ &\le \int_{\tilde{B}} F(\exp(\beta K_\mu(w)))\, dm(w) + \int_{\tilde{B}} G(J_\Phi(w))\, dm(w)\\ &\le \int_{\tilde{B}} \exp((\exp(\beta K_\mu(w)))^{1/p})\, dm(w) + \int_{\tilde{B}} C_p J_\Phi(w)(\log(1 + C_p J_\Phi(w)))^p\, dm(w).\end{aligned} \tag{4-25}$$

We apply this by using $p > \beta/q$, so that

$$\big(\exp(\beta K_\mu(w))\big)^{1/p} \le \exp(q K_\mu(w)).$$

Thus

$$\int_{\tilde{B}} \exp\big(\big(\exp(\beta K_\mu(w))\big)^{1/p}\big)\, dm(w) \le \int_{\tilde{B}} \exp\big(\exp(q K_\mu(w))\big)\, dm(w) < \infty.$$

The last term in (4-25) is finite by (4-15), and thus the claim follows. □

4B. ***Asymptotics of the phase function of the exponentially growing solution.*** Let $\mu \in B^p_{\exp}(B(R_0))$, $k \in \mathbb{C} \setminus \{0\}$ and $\lambda \in \mathbb{C}$ satisfy $|\lambda| \le 1$. Then using Lemmas 3.3 and 3.4, with the affine weight $\mathcal{A}(t) = pt - p$ corresponding to the gauge function Q, we see that the equation

$$\bar\partial_z f_k(z) = \lambda \mu(z)\, \overline{\partial_z f_k(z)} \qquad \text{for a.e. } z \in \mathbb{C}, \tag{4-26}$$

$$f_k(z) = e^{ikz}\Big(1 + O\Big(\frac{1}{z}\Big)\Big) \quad \text{as } |z| \to \infty, \tag{4-27}$$

has the unique solution $f_k \in W^{1,Q}_{\mathrm{loc}}(\mathbb{C})$. Moreover, this solution can be written in the form

$$f_k(z) = e^{ik\varphi_k(z)}, \tag{4-28}$$

where $\varphi_k : \mathbb{C} \to \mathbb{C}$ is a homeomorphism satisfying

$$\bar\partial \varphi_k(z) = -\frac{\lambda \bar k}{k} \mu(z) e_{-k}(\varphi_k(z))\, \overline{\partial \varphi_k(z)} \quad \text{for a.e. } z \in \mathbb{C}, \tag{4-29}$$

$$\varphi_k(z) = z + \mathcal{O}\Big(\frac{1}{z}\Big) \quad \text{as } |z| \to \infty. \tag{4-30}$$

Below, we set $f_k(z) = f(z,k)$ and $\varphi_k(z) = \varphi(z,k)$ and estimate the next functions φ_k in the Orlicz space $X^q(\mathbb{C})$. The following lemma is a generalization of results of [Astala and Päivärinta 2006] to the Orlicz space setting.

Lemma 4.6. *Assume that $\nu \in B^p_{\exp}(B(R_0))$ for all $0 < p < \infty$. For $k \in \mathbb{C} \setminus \{0\}$, let $\Phi_k \in W^{1,1}(\mathbb{C})$ be the solution of*

$$\bar\partial \Phi_k(z) = -\frac{\bar k}{k} \nu(z) e_{-k}(z)\, \partial \Phi_k(z) \quad \textit{for a.e. } z \in \mathbb{C}, \tag{4-31}$$

$$\Phi_k(z) = z + \mathcal{O}\Big(\frac{1}{z}\Big). \tag{4-32}$$

Then for all $\varepsilon > 0$, there exists $C_0 > 0$ such that $\bar\partial_z \Phi_k(z) = g_k(z) + h_k(z)$, where $g_k, h_k \in X^q(\mathbb{C})$ are supported in $B(R_0)$ and

$$\sup_{k \in \mathbb{C}\setminus\{0\}} \|h_k\|_{X^q} < \varepsilon, \tag{4-33}$$

$$\sup_{k \in \mathbb{C}\setminus\{0\}} \|g_k\|_{X^q} < C_0, \tag{4-34}$$

$$\lim_{k\to\infty} \hat g_k(\xi) = 0, \tag{4-35}$$

where for all compact sets $S \subset \mathbb{C}$, the convergence in (4-35) *is uniform for $\xi \in S$.*

Proof. Let us define

$$\tilde\nu_k(z) = -\bar k k^{-1} \nu(z)$$

for $k \in \mathbb{C} \setminus \{0\}$. Note that then for any $p > 0$, there is $N > 0$ such that $\tilde\nu_k(\,\cdot\,, k) e_{-k}(\,\cdot\,) \in B^p_{\exp,N}(B(R_0))$ for all $k \in \mathbb{C} \setminus \{0\}$. By Theorem 4.2,

$$\lim_{n\to\infty} \Big\| \bar\partial \Phi_k - \sum_{n=0}^{\infty} (\tilde\nu_k e_{-k} S)^n (\tilde\nu_k e_{-k}) \Big\|_{X^q(\mathbb{C})} = 0$$

uniformly in $k \in \mathbb{C} \setminus \{0\}$. For $m \in \mathbb{Z}_+$, we define

$$g_k(z) = g_k^{(m)}(z) = -\sum_{n=0}^{m} (\tilde\nu_k e_{-k} S)^n (\tilde\nu_k e_{-k}),$$

$$h_k(z) = h_k^{(m)}(z) = -\sum_{n=m+1}^{\infty} (\tilde\nu_k e_{-k} S)^n (\tilde\nu_k e_{-k}).$$

For given $\varepsilon > 0$, we can choose m so large that (4-33) holds for all $k \in \mathbb{C} \setminus \{0\}$, and then using Lemma 4.3, choose C_0 so that (4-34) holds for all $k \in \mathbb{C} \setminus \{0\}$.

Next, we show (4-35) when ε and m are fixed so that (4-33) and (4-34) hold. We can write

$$g_k(z) = -\sum_{n=0}^{m} e_{-nk} G_n, \quad G_n = \left(\frac{\bar{k}}{k}\right)^{n+1} \nu S_n(k)\nu \cdots \nu S_1(k)\nu,$$

where $S_j(k)$ is the Fourier multiplier

$$(S_j(k)\phi)^\wedge(\xi) = \boldsymbol{m}(\xi + jk)\hat{\phi}(\xi), \quad \boldsymbol{m}(\xi) = \frac{\bar{\xi}}{\xi}.$$

The proof of [Astala and Päivärinta 2006, Lemma 7.3] for $n \geq 1$ and the Riemann–Lebesgue lemma for $n = 0$ yield that for any $\tilde{\varepsilon} > 0$, there exists $R(n, \tilde{\varepsilon}) \geq 0$ such that, for $n \leq m$,

$$|\widehat{G}_n(\xi)| \leq (n+1)\kappa^n \tilde{\varepsilon} \quad \text{for } |\xi| > R(n, \tilde{\varepsilon}),$$

where $\kappa = \|\nu\|_{L^\infty} \leq 1$. Thus for $n \leq m$,

$$|\widehat{G}_n(\xi)| \leq (m+1)\tilde{\varepsilon} \quad \text{for } |\xi| > R_0 = \max_{n \leq m} R(n, \tilde{\varepsilon}),\ n = 0, 1, 2, \ldots, m. \tag{4-36}$$

As

$$(e_{-nk} G_n)^\wedge(\xi) = \widehat{G}_n(\xi - nk),$$

we see that for any $L > 0$, there is $k_0 > 0$ such that if $|k| > k_0$ then $j|k| - L > R_0$ for $1 \leq n \leq m$. Then it follows from (4-36) that if $|k| > k_0$, then

$$\sup_{|\xi| < L} |\hat{g}_k(\xi)| \leq (m+1)^2 \tilde{\varepsilon}.$$

This proves the limit (4-35), with the convergence being uniform for ξ belonging in a compact set. □

Proposition 4.7. *Assume that $\nu \in B^p_{\exp}(B(R_0))$ with $p > 4$ and $\Phi_k(z)$ is the solution of* (4-31)–(4-32). *Then*

$$\lim_{k \to \infty} \Phi_k(z) = z \quad \textit{uniformly for } z \in \mathbb{C}. \tag{4-37}$$

Proof. Step 1: We will first show that for all q with $4 < q < p$, we have $\bar{\partial}_z \Phi_k(z) \to 0$ weakly in $X^q(\mathbb{C})$ as $k \to \infty$. Let $\eta \in X^{-q}(\mathbb{C})$ and $\varepsilon_1 > 0$. By Theorem 4.2, there is $C_1 > 0$ such that

$$\sup_k \|\bar{\partial}\Phi_k\|_{X^q} \leq C_1.$$

Since $C_0^\infty(\mathbb{C})$ is dense in $X^{-q}(\mathbb{C})$ (see [Krasnosel'skiĭ and Rutickiĭ 1961, Section II.10]), we can find a function $\eta_0 \in C_0^\infty(\mathbb{C})$ such that

$$\|\eta - \eta_0\|_{X^{-q}} \leq \min(1, \varepsilon_1 / C_1).$$

Then

$$|\langle \eta, \bar{\partial}\Phi_k \rangle| \leq |\langle \eta_0, \bar{\partial}\Phi_k \rangle| + \|\eta - \eta_0\|_{X^{-q}(\mathbb{C})} \|\bar{\partial}\Phi_k\|_{X^q(\mathbb{C})}, \tag{4-38}$$

where the second term on the right-hand side is smaller than ε_1. Moreover, by Lemma 4.6, we can write

$$\bar{\partial}\Phi_k = h_k + g_k$$

so that (4-33)–(4-35) are satisfied for $\varepsilon = \varepsilon_1(\|\eta\|_{X^{-q}} + 1)^{-1}$ and some $C_0 > 0$. Then $|\langle \eta_0, h_k \rangle| \leq \varepsilon_1$.

Since $\hat{\eta}_0$ is a rapidly decreasing function, $\hat{g}_k(\xi)$ is uniformly bounded for $\xi \in \mathbb{C}$ and $k \in \mathbb{C} \setminus \{0\}$ by Lemma 4.6, and $\hat{g}_k \to 0$ uniformly in all bounded domains as $k \to \infty$, we see that

$$\langle \eta_0, g_k \rangle = \langle \hat{\eta}_0, \hat{g}_k \rangle \to 0 \quad \text{as } k \to \infty. \tag{4-39}$$

Combining these, we see that $\langle \eta_0, \bar{\partial}\Phi_k \rangle \to 0$ as $k \to \infty$, and thus $\bar{\partial}_z \Phi_k(z) \to 0$ weakly in $X^q(\mathbb{C})$ as $k \to \infty$.

Step 2: Next we show the pointwise convergence

$$\lim_{k \to \infty} \bar{\partial}_z \Phi_k(z) = 0. \tag{4-40}$$

To this end, we observe that the function

$$\eta_z(w) = \frac{1}{\pi(w-z)} \chi_{B(R_0)}(w)$$

satisfies $\eta_z \in X^{-q}(\mathbb{C})$ for $q > 1$. Since $\Phi_k(z) - z = \mathcal{O}(1/z)$ and $\bar{\partial}\Phi_k$ is supported in $\overline{B(R_0)}$, we have

$$\Phi_k(z) = z - \frac{1}{\pi} \int_{B(R_0)} (w-z)^{-1}\, \bar{\partial}_w \Phi_k(w)\, dm(w) = z - \langle \eta_z, \bar{\partial}\Phi_k \rangle. \tag{4-41}$$

As $\bar{\partial}\Phi_k \to 0$ weakly in $X^q(\mathbb{C})$, we see (4-40) holds for all $z \in \mathbb{C}$.

Step 3: By (3-14) and (3-17), we see that the family $\{\Phi_k(z)\}_{k \in \mathbb{C}\setminus\{0\}}$ of homeomorphisms has a uniform modulus of continuity in compact sets. Moreover, since

$$\sup_k \|\bar{\partial}\Phi_k\|_{L^1(\mathbb{C})} \le \sup_k \|\bar{\partial}\Phi_k\|_{X^q(B(R_0))} = C_2 < \infty,$$

we obtain by (4-40), for $|z| > R_0 + 1$, that

$$|\Phi_k(z) - z| = |\langle \eta_z, \bar{\partial}\Phi_k \rangle| \le \frac{C}{|z|} \|\bar{\partial}\Phi_k\|_{L^1(\mathbb{C})} \le \frac{CC_2}{|z|}. \tag{4-42}$$

Thus, as the functions $\{\Phi_k(z)\}_{k \in \mathbb{C}\setminus\{0\}}$ are uniformly equicontinuous in compact sets, (4-42) and the pointwise convergence (4-40) yield the uniform convergence (4-37). □

4C. *Properties of the solutions of the nonlinear Beltrami equation.* Let $\lambda \in \mathbb{C}$, $|\lambda| \le 1$ and $\mu(z)$ be supported in $\overline{B(R_0)}$, $R_0 \ge 1$, and assume that $K = K_\mu$ satisfies (4-1) with q, $C_0 > 0$ and $R_1 = 1$. Motivated by Lemma 3.4, we consider next the solutions φ_k of the equation

$$\bar{\partial}_z \varphi_\lambda(z,k) = -\lambda \frac{\bar{k}}{k} \mu(z) e_{-k}(\varphi_\lambda(z,k))\, \overline{\partial_z \varphi_\lambda(z,k)}, \quad z \in \mathbb{C}, \tag{4-43}$$

$$\varphi_\lambda(z,k) = z + O\Big(\frac{1}{z}\Big) \quad \text{as } |z| \to \infty. \tag{4-44}$$

Let $\psi_\lambda(\,\cdot\,,k) = \varphi_\lambda(\,\cdot\,,k)^{-1}$ be the inverse function of $\varphi_\lambda(\,\cdot\,,k)$. A simple computation based on differentiation of the identity $\psi_\lambda(\varphi_\lambda(z,k),k) = z$ in the z-variable shows that

$$\bar{\partial}_z \psi_\lambda(z,k) = -\lambda \frac{\bar{k}}{k} \mu(\psi_\lambda(z,k)) e_{-k}(z)\, \partial_z \psi_\lambda(z,k), \quad z \in \mathbb{C}, \tag{4-45}$$

$$\psi_\lambda(z,k) = z + O\Big(\frac{1}{z}\Big) \quad \text{as } |z| \to \infty. \tag{4-46}$$

Define

$$\nu(z) = -\lambda \frac{\bar{k}}{k} \mu(z)$$

and consider the equations (4-43) and (4-45) simultaneously by defining the sets

$$B_\mu = \{(\varphi, \nu) : |\nu| \le |\mu| \text{ a.e. and } \varphi : \mathbb{C} \to \mathbb{C} \text{ is a homeomorphism with } \bar\partial\varphi = \nu\overline{\partial\varphi},\ \varphi(z) = z + O(z^{-1})\}$$

and, for Q defined by (1-8),

$$\mathcal{G}_\mu = \{g \in W^{1,Q}_{\mathrm{loc}}(\mathbb{C}) : \bar\partial g = (\nu \circ \varphi^{-1})\partial g,\ g(z) = z + O(z^{-1}),\ (\varphi, \nu) \in B_\mu\}.$$

Now $\exp(\exp(qK_\mu)) \in L^1(B(R_0))$ with some $0 < q < \infty$ and $|\nu| \le |\mu|$ almost everywhere. Then $K_\nu(z) \le K_\mu(z)$ almost everywhere. Let

$$\bar\partial\varphi = \nu\, \overline{\partial\varphi} \quad \text{in } \mathbb{C}, \qquad \varphi(z) = z + O(z^{-1}),$$

so that $\bar\partial\varphi = \tilde\nu\, \partial\varphi$ with $|\tilde\nu(z)| = |\nu(z)|$. Then for $\psi = \varphi^{-1}$, we have $K(z, \psi) = K_\nu(\psi(z))$; see (2-14). Thus by Lemma 4.5, we have

$$\sup_{g \in \mathcal{G}_\mu} \|\exp(\beta K(\,\cdot\,, g))\|_{L^1(B(R))} = \sup_{(\varphi,\nu)\in B_\mu} \|\exp(\beta K_\nu \circ \varphi^{-1})\|_{L^1(B(R))} < \infty \tag{4-47}$$

for all $\beta > 0$ and $R > 0$. Using this and Theorem 2.2, we see that the functions $g \in \mathcal{G}_\mu$ are homeomorphisms. Moreover, recall that $\mu \in B^p_{\exp}(B(R))$ for all $p \in (1, \infty)$. Thus for $g \in \mathcal{G}_\mu$, the condition $g \in W^{1,Q}_{\mathrm{loc}}(\mathbb{C})$ is equivalent to (see (4-7) and (4-8)) $Dg \in X^{-1}_{\mathrm{loc}}(\mathbb{C})$. Furthermore by (4-13), we have

$$\sup_{(\varphi,\nu)\in B_\mu} \|J_\varphi\|_{X^{1,q}(B(R))} < \infty \tag{4-48}$$

for all $q > 0$.

Lemma 4.8. *The set $\mathcal{G}_\mu$ is relatively compact in the topology of uniform convergence.*

Proof. Let $(\varphi, \nu) \in B_\mu$ and $\psi = \varphi^{-1}$ and

$$\bar\partial g = (\nu \circ \varphi^{-1})\partial g, \quad g(z) = z + O(z^{-1}).$$

As μ is supported in $B(R_0)$, the function φ is analytic outside $\overline{B(R_0)}$; we see using (3-7) for the function φ that for $R > 0$, we have $\varphi(B(R)) \subset B(R + 3R_0)$, $\psi(B(R)) \subset B(R + 3R_0)$, and that ψ is analytic outside $\overline{B(4R_0)}$.

Thus (3-7) and the same arguments which we used to prove the estimate (3-27) yield that for $R > 0$,

$$\begin{aligned}\|Q(|Dg|)\|_{L^1(B(R))} &\le \pi(R + 3R_0)^2 + \int_{B(R)} \exp\bigl(qK_\nu(\psi(w)) - q\bigr)\, dm(w)\\ &\le \pi(R + 3R_0)^2 + \int_{B(R+3R_0)} \exp(qK_\nu(z) - q) J_\varphi(z)\, dm(z),\end{aligned} \tag{4-49}$$

where $Q(t) = |t|^2/\log(|t| + e)$. We will next use Young's inequality (A-7) with the admissible pair (F, G), where (see [Krasnosel'skiĭ and Rutickiĭ 1961, Chapter 1.3])

$$F(t) = e^t - t - 1, \quad G(t) = (1 + t)\log(1 + t) - t. \tag{4-50}$$

By Young's inequality, we have

$$\int_{B(R+3R_0)} \exp(qK_\nu(z)-q)J_\varphi(z)\,dm(z) \le \int_{B(R+3R_0)} \exp\bigl(\exp(qK_\nu(w)-q)\bigr)\,dm(w) + \int_{B(R+3R_0)} (1+J_\varphi(w))\log(1+J_\varphi(w))\,dm(w).$$

This, (4-1), (4-48), and (4-49) show that there is a constant $C(R,\mu)$ such that for $g \in \mathcal{G}_\mu$,

$$\|Q(|Dg|)\|_{L^1(B(R))} \le C(R,\mu). \tag{4-51}$$

As $g \in \mathcal{G}_\mu$ are homeomorphisms, this, (3-14) and (A-1)–(A-3) imply that functions $g \in \mathcal{G}_\mu$ are equicontinuous in compact sets of $\mathbb{C}$. As $\operatorname{supp}(\nu \circ \psi) \subset B(4R_0)$, functions $g \in \mathcal{G}_\mu$ are analytic outside the disc $B(4R_0)$ and the inequality (3-7) yields, for $R > 0$ and $g \in \mathcal{G}_\mu$, that

$$g(B(R)) \subset B(R+12R_0).$$

By the Arzelà–Ascoli theorem, these imply that the set $\{g|_{B(R)} : g \in \mathcal{G}_\mu\}$ is relatively compact in the topology of uniform convergence in $B(R)$ for any $R > 0$. Thus by using a diagonalization argument, we see that for an arbitrary sequence $g_n \in \mathcal{G}_\mu$, $n = 1, 2, \ldots$, there is a subsequence g_{n_j} which converges uniformly in all discs $B(R)$, $R > 0$. Finally, by Young's inequality (see the Appendix), we get using the same notations as in (4-41) that for $|z| > 4R_0 + 1$,

$$\begin{aligned} |g_k(z) - z| &= \left| \frac{1}{\pi} \int_{B(4R_0)} (w-z)^{-1}\, \bar\partial_w g_k(w)\,dm(w) \right| \\ &\le \frac{1}{\pi(|z|-4R_0)} \int_{B(4R_0)} \bigl(Q(|\bar\partial_w g_k(w)|) + G_0(1)\bigr)\,dm(w), \end{aligned} \tag{4-52}$$

where $Q(t)$ and $G_0(t) = |t|^2 \log(|t|+1)$ form a Young complementary pair. Thus

$$|g_k(z) - z| \le \frac{C_\mu}{|z| - 4R_0} \quad \text{for } |z| > 4R_0 + 1.$$

Using this and the uniform convergence of g_{n_j} in all discs $B(R)$, $R > 0$, we see that g_n has a subsequence converging uniformly in $\mathbb{C}$. □

Theorem 4.9. *Let* $\lambda, k \in \mathbb{C} \setminus \{0\}$, $|\lambda| = 1$. *Assume that* $\varphi_\lambda(z,k)$ *satisfies* (4-43)–(4-44) *with* μ *supported in* $\mathbb{D}$ *which satisfies* (4-1) *with* $q > 0$ *and* $R_1 = 1$. *Then*

$$\lim_{k\to\infty} \varphi_\lambda(z,k) = z$$

uniformly in $z \in \mathbb{C}$ *and* $|\lambda| = 1$.

Proof. Let $\psi_\lambda(\,\cdot\,,k)$ be the inverse function of $\varphi_\lambda(\,\cdot\,,k)$. It is sufficient to show that

$$\lim_{k\to\infty} \psi_\lambda(z,k) = z$$

uniformly in $z \in \mathbb{C}$ and $|\lambda| = 1$.

Then, $\psi_\lambda(\,\cdot\,,k)$ is the solution of (4-45)–(4-46). Define

$$\nu(z) = -\lambda \bar{k} k^{-1} \mu(z)$$

and note that $|\nu(z)| = |\mu(z)|$. Hence

$$\big(\varphi_\lambda(\,\cdot\,,k),\, \nu(\,\cdot\,)\, e_{-k}(\,\cdot\,)\big) \in B_\mu$$

and $\psi_\lambda(\,\cdot\,,k) \in \mathcal{G}_\mu$. Moreover, as $\varphi_\lambda(\,\cdot\,,k)$ is homeomorphism in $\mathbb{C}$ and analytic outside of $B(1)$, it follows from (3-7) with $r = 1$ that $\varphi_\lambda(\,\cdot\,,k)$ maps the ball $B(1)$ into $B(4)$ and moreover, its inverse $\psi_\lambda(\,\cdot\,,k)$ maps the disc $B(4)$ into $B(5)$ and $\mathbb{C} \setminus B(4)$ into $\mathbb{C} \setminus B(1)$.

It follows from Lemma 4.8 that if the claim is not valid, there are sequences $(\lambda_n)_{n=1}^\infty$, $|\lambda_n| = 1$, and $(k_n)_{n=1}^\infty$, $k_n \to \infty$, such that

$$\psi_\infty(z) = \lim_{n\to\infty} \psi_{\lambda_n}(z, k_n), \tag{4-53}$$

where the convergence is uniform, $z \in \mathbb{C}$, and $\psi_\infty(z)$ is not equal to z. Thus, to prove the claim, it is enough to show that any limit of form (4-53) satisfies $\psi_\infty(z) = z$. Note that by considering subsequences, we can assume that $\lambda_n \to \lambda$ and $\bar{k}_n k_n^{-1} \to \beta$ as $n \to \infty$, where $|\lambda| = |\beta| = 1$. Next define $\nu_0(z) = -\lambda\beta\mu(z)$.

Let us consider the solution of

$$\bar\partial_z \Phi_\lambda(z,k) = \nu_0(\psi_\infty(z))\, e_{-k}(z)\, \partial_z \Phi_\lambda(z,k), \tag{4-54}$$

$$\Phi_\lambda(z,k) = z + O\Big(\frac{1}{z}\Big) \quad \text{as } |z| \to \infty. \tag{4-55}$$

We note that here $\nu_0(\psi_\infty(z)) = 0$ for $|z| > 4$ as ν_0 is supported in $\bar{B}(1)$ and ψ_∞ maps $\mathbb{C} \setminus \bar{B}(4)$ into $\mathbb{C} \setminus \bar{B}(1)$. By Proposition 4.7, $\Phi_\lambda(z,k) \to z$ as $k \to \infty$ uniformly in $z \in \mathbb{C}$. Since for every $z \in \mathbb{C}$, the function $\eta_z : w \mapsto \chi_{B(4)}(w)(z-w)^{-1}$ is in $X^{-q}(\mathbb{C})$ for $q > 1$, we obtain, using (4-41), that

$$\begin{aligned} |\psi_{\lambda_n}(z,k_n) - \Phi_\lambda(z,k_n)| &= \frac{1}{\pi}\left| \int_{B(4)} (w-z)^{-1}\, \bar\partial_w\big(\psi_{\lambda_n}(w,k_n) - \Phi_\lambda(w,k_n)\big)\, dm(w) \right| \\ &\le \|\eta_z\|_{X^{-q}} \big\| \bar\partial(\psi_{\lambda_n}(\,\cdot\,,k_n) - \Phi_\lambda(\,\cdot\,,k_n)) \big\|_{X^q(B(4))}. \end{aligned} \tag{4-56}$$

Let us next assume that we can prove that

$$\lim_{n\to\infty} \big\| \mu \circ \psi_{\lambda_n}(\,\cdot\,,k_n) - \mu \circ \psi_\infty(\,\cdot\,,k_n) \big\|_{L^s(\mathbb{C})} = 0 \quad \text{for some } s > 2. \tag{4-57}$$

If this is the case, let $p \in (4q, \infty)$. By assumption (4-1) and Lemma 4.5, there is N such that the Beltrami coefficients of functions $\psi_{\lambda_n}(\,\cdot\,,k_n)$ are in $B^p_{\exp,N}(\mathbb{D})$ for all $n \in \mathbb{Z}_+$ and $p > 4$. By Theorem 4.2 and (4-57),

$$\lim_{n\to\infty} \big\| \bar\partial\big(\psi_{\lambda_n}(\,\cdot\,,k_n) - \Phi_\lambda(\,\cdot\,,k_n)\big) \big\|_{X^q(\mathbb{C})} = 0.$$

As $\lim_{n\to\infty} \Phi_\lambda(z,k_n) = z$ uniformly in $z \in \mathbb{C}$, this and (4-56) show that $\psi_\infty(z) = z$.

Thus, to prove the claim it is enough to show (4-57). First, as $\psi_{\lambda_n}(\,\cdot\,,k_n) \to \psi_\infty(\,\cdot\,)$ uniformly as $n \to \infty$ and as $\psi_{\lambda_n}(\,\cdot\,,k_n)$ maps $\mathbb{C} \setminus B(3)$ into $\mathbb{C} \setminus B(2)$, we see using the dominated convergence theorem that the formula (4-57) is valid when μ is replaced by a smooth compactly supported function. Next, let (F, G) be the complementary Young pair given by (4-50) and $E_F(B(R))$ be the closure of $L^\infty(B(R))$ in $X_F(B(R))$. By [Adams 1975, Theorem 8.21], the set $C_0^\infty(\mathbb{D})$ is dense in $E_F(\mathbb{D})$ with respect to the norm of X_F. Thus when μ is a nonsmooth Beltrami coefficient satisfying the assumption (4-1) and $\varepsilon > 0$, we can find a smooth function $\theta \in C_0^\infty(\mathbb{D})$, $\|\theta\|_\infty < 2$ such that $\|\mu-\theta\|_F < \varepsilon$. Then, since $|\mu-\theta|$ is supported in $\mathbb{D}$ and bounded by 3, we have

$$\begin{aligned}\big\|\mu\circ\psi_{\lambda_n}(\cdot,k_n)-\theta\circ\psi_{\lambda_n}(\cdot,k_n)\big\|^s_{L^s(\mathbb{C})} &= \int_{\mathbb{D}} |\mu(z)-\theta(z)|^s J_{g_n}(z)\,dm(z)\\ &\le 3^{s-1}\bigg(\int_{\mathbb{D}} F(|\mu(z)-\theta(z)|)\,dm(z)\bigg)\bigg(\int_{\mathbb{D}} G(J_{g_n}(z))\,dm(z)\bigg), \quad (4\text{-}58)\end{aligned}$$

where g_n is the inverse of the function $\psi_{\lambda_n}(\cdot,k_n)$. Then,

$$\int_{\mathbb{D}} G(J_{g_n})\,dm \le C\|J_{g_n}\|_{X^{1,1}(B(2))}$$

and by (4-48), $\|J_{g_n}\|_{X^{1,1}(\mathbb{D})}$ is uniformly bounded in n. Using (4-58) and (A-5), we see that (4-57) holds for all μ satisfying the assumption (4-1) and thus claim of the theorem follows. □

4D. $\bar\partial$-*equations in k-planes.* Let us consider a Beltrami coefficient $\mu \in B^p_{\exp}(\mathbb{D})$ and approximate μ with functions μ_n supported in $\mathbb{D}$ for which

$$\lim_{n\to\infty} \mu_n(z) = \mu(z) \quad \text{and} \quad \|\mu_n\|_\infty \le c_n < 1;$$

see, e.g., (3-19). Let $f_\mu(\cdot,k) \in W^{1,Q}_{\mathrm{loc}}(\mathbb{C})$ be the solution of the equations

$$\bar\partial_z f_\mu(z,k) = \mu(z)\,\overline{\partial_z f_\mu(z,k)} \qquad \text{for a.e. } z \in \mathbb{C}, \quad (4\text{-}59)$$

$$f_\mu(z,k) = e^{ikz}\Big(1+\mathcal{O}_k\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z|\to\infty, \quad (4\text{-}60)$$

and $f_{\mu_n}(\cdot,k) \in W^{1,Q}_{\mathrm{loc}}(\mathbb{C})$ be the solution of the similar equations of Beltrami coefficients μ_n and μ; see Lemma 3.4. Here $\mathcal{O}_k(h(z))$ means a function of (z,k) that satisfies $|\mathcal{O}_k(h(z))| \le C(k)|h(z)|$ for all z with some constant $C(k)$ depending on $k \in \mathbb{C}$. Let

$$\varphi_\mu(z,k) = (ik)^{-1}\log(f_\mu(z,k)), \quad \varphi_{\mu_n}(z,k) = (ik)^{-1}\log(f_n(z,k));$$

see (3-5). Then by (3-7), we have

$$|\varphi_{\mu_n}(z,k)| \le |z|+3, \quad |\varphi_\mu(z,k)| \le |z|+3. \quad (4\text{-}61)$$

By the proof of Lemma 3.4, we see that by choosing a subsequence of μ_n, $n \in \mathbb{Z}_+$, which we continue to denote by μ_n, we can assume that

$$\lim_{n\to\infty} \varphi_{\mu_n}(z,k) = \varphi_\mu(z,k) \quad \text{uniformly in } (z,k) \in B(R)\times\{k_0\} \text{ for all } R>0 \text{ and } k_0 \in \mathbb{C}. \quad (4\text{-}62)$$

Let us write the solutions f_{μ_n} and f_μ as

$$\begin{aligned} f_{\mu_n}(z,k) &= e^{ik\varphi_{\mu_n}(z,k)} = e^{ikz}M_{\mu_n}(z,k),\\ f_\mu(z,k) &= e^{ik\varphi_\mu(z,k)} = e^{ikz}M_\mu(z,k).\end{aligned}$$

Similar notations are introduced when μ is replaced by $-\mu$ etc. Let

$$h^{(+)}_{\mu_n}(z,k)=\tfrac{1}{2}\big(f_{\mu_n}(z,k)+f_{-\mu_n}(z,k)\big),$$
$$h^{(-)}_{\mu_n}(z,k)=\tfrac{1}{2}i\big(\overline{f_{\mu_n}(z,k)}-\overline{f_{-\mu_n}(z,k)}\big),$$

and

$$u^{(1)}_{\mu_n}(z,k)=h^{(+)}_{\mu_n}(z,k)-ih^{(-)}_{\mu_n}(z,k),$$
$$u^{(2)}_{\mu_n}(z,k)=-h^{(-)}_{\mu_n}(z,k)+ih^{(+)}_{\mu_n}(z,k).$$

Then by (4-61), $h^{(+)}_{\mu_n}(z,k)$ and $h^{(-)}_{\mu_n}(z,k)$ are uniformly bounded for $(z,k)\in B(R_1)\times B(R_2)$ for any $R_1, R_2>0$. By (4-62), we can define the pointwise limits

$$\lim_{n\to\infty} h^{(\pm)}_{\mu_n}(z,k)=h^{(\pm)}_{\mu}(z,k),\qquad \lim_{n\to\infty} u^{(j)}_{\mu_n}(z,k)=u^{(j)}_{\mu}(z,k),\qquad j=1,2. \tag{4-63}$$

The above formulae imply

$$u^{(2)}_{\mu}(z,k)=iu^{(1)}_{-\mu}(z,k)\quad\text{and}\quad u^{(1)}_{\mu}(z,k)=-iu^{(2)}_{-\mu}(z,k). \tag{4-64}$$

Moreover, for

$$\tau_{\mu_n}(k)=\tfrac{1}{2}\big(\overline{t_{\mu_n}(k)}-\overline{t_{-\mu_n}(k)}\big),\qquad \tau_{\mu_n}(k)=\tfrac{1}{2}\big(\overline{t_{\mu_n}(k)}-\overline{t_{-\mu_n}(k)}\big),$$

and

$$t_{\pm\mu_n}(k)=\frac{i}{2\pi}\int_{\partial\mathbb{D}} M_{\pm\mu_n}(z,k)\,dz,\qquad t_{\pm\mu}(k)=\frac{i}{2\pi}\int_{\partial\mathbb{D}} M_{\pm\mu}(z,k)\,dz,$$

we see using the dominated convergence theorem that $\lim_{n\to\infty} t_{\mu_n}(k)=t_\mu(k)$ for all $k\in\mathbb{C}$, and hence

$$\lim_{n\to\infty}\tau_{\mu_n}(k)=\tau_\mu(k)\quad\text{for all } k\in\mathbb{C}. \tag{4-65}$$

Then, as $|\mu_n|\le c_n<1$ correspond to conductivities σ_n satisfying $\sigma_n,\sigma_n^{-1}\in L^\infty(\mathbb{D})$, we have by [Astala and Päivärinta 2006, Formula (8.2)] the $\bar\partial$-equations with respect to the k-variables,

$$\bar\partial_k u^{(j)}_{\mu_n}(z,k)=-i\tau_{\mu_n}(k)\overline{u^{(j)}_{\mu_n}(z,k)},\quad k\in\mathbb{C},\ j=1,2; \tag{4-66}$$

see also [Nachman 1988; 1996] for a different formulation of such equations. For $z\in\mathbb{C}$, functions $u^{(j)}_{\mu_n}(z,\cdot)$, $n\in\mathbb{Z}_+$, are uniformly bounded in $B(R)$ for all $R>0$; the limit (4-63) and the dominated convergence theorem imply that $u^{(j)}_{\mu_n}(z,\cdot)\to u^{(j)}_{\mu}(z,\cdot)$ as $n\to\infty$ in $L^p(B(R))$ for all $p<\infty$ and $R>0$. Since the functions $|\tau_{\mu_n}(k)|$, $n\in\mathbb{Z}_+$, are uniformly bounded in compact sets, the pointwise limits (4-63), (4-65) and the equation (4-66) yield that

$$\bar\partial_k u^{(j)}_{\mu}(z,k)=-i\tau_{\mu}(k)\overline{u^{(j)}_{\mu}(z)},\quad k\in\mathbb{C},\ j=1,2, \tag{4-67}$$

holds for all $z\in\mathbb{C}$ in the sense of distributions and $u^{(j)}_{\mu}(z,\cdot)\in W^{1,p}_{\mathrm{loc}}(\mathbb{C})$ for all $p<\infty$.

4E. *Proof of uniqueness results for isotropic conductivities.*

Proof of Theorem 1.9. Let us consider isotropic conductivities σ_j, $j=1,2$. Due to the above proven results, the proof will go along the lines of Section 8 of [Astala and Päivärinta 2006], where L^∞-conductivities are considered, and its reformulation, presented in Section 18 of [Astala et al. 2009] in a quite straightforward way, when the changes explained below are made. The key point is the following proposition.

Proposition 4.10. *Assume that $\mu \in B^p_{\exp}(\mathbb{D})$ and let $f_{\pm\mu}(z,k)$ satisfy* (4-59)–(4-60) *with the Beltrami coefficients $\pm\mu$. Then $f_{\pm\mu}(z,k) = e^{izk} M_{\pm\mu}(z,k)$, where*

$$\operatorname{Re} \frac{M_{+\mu}(z,k)}{M_{-\mu}(z,k)} > 0 \tag{4-68}$$

for every $z, k \in \mathbb{C}$.

Proof. Let us consider the Beltrami coefficients $\mu_n(z)$, $n \in \mathbb{Z}_+$, defined in Section 4D that converge pointwise to $\mu(z)$ and satisfy $|\mu_n| \le c_n < 1$. By Lemma 3.2, the functions $M_{\pm\mu_n}(z,k)$ do not attain the value zero anywhere. By [Astala and Päivärinta 2006, Proposition 4.3], the inequality (4-68) holds for the functions $M_{\pm\mu_n}(z,k)$. Then, $f_{\pm\mu_n}(z,k) \to f_{\pm\mu}(z,k)$ as $n \to \infty$ for all $k, z \in \mathbb{C}$, and thus we see that

$$\operatorname{Re} \frac{M_{+\mu}(z,k)}{M_{-\mu}(z,k)} = \lim_{n\to\infty} \operatorname{Re} \frac{M_{+\mu_n}(z,k)}{M_{-\mu_n}(z,k)} \ge 0. \tag{4-69}$$

To show that the equality does not hold in (4-69), we assume the opposite. In this case, there are z_0 and k_0 such that

$$M_{+\mu}(z_0,k_0) = it M_{-\mu}(z_0,k_0) \tag{4-70}$$

for some $t \in \mathbb{R} \setminus \{0\}$. Then

$$f(z,k_0) = e^{ik_0 z}\big(M_{+\mu}(z,k_0) - it M_{-\mu}(z,k_0)\big)$$

is a solution of (4-59) and satisfies the asymptotics

$$f(z,k_0) = (1-it) e^{ik_0 z}\Big(1 + \mathcal{O}\Big(\frac{1}{z}\Big)\Big) \quad \text{for } |z| \to \infty.$$

By using (2-13) to write (4-59) in the form (3-13) and applying Lemma 3.2, we see that the solution $f(z,k_0)$ can be written in the form

$$f(z,k_0) = (1-it) e^{ik_0 \varphi(z)}.$$

This is in contradiction with (4-70), which would be implied by $f(z_0,k_0)=0$, and thus proves (4-68). □

Let $f_{\pm\mu}(z,k)$ be as in Proposition 4.10 and use below for the functions defined in (4-63) the shorthand notation $u^{(1)}_\mu(z,k) = u_1(z,k)$ and $u^{(2)}_\mu(z,k) = u_2(z,k)$. Then $u_1(z,k)$ and $u_2(z,k)$ are solutions of (4-67). A direct computation shows also that

$$\nabla\cdot\sigma\nabla u_1(\,\cdot\,,k) = 0 \quad \text{and} \quad \nabla\cdot\frac{1}{\sigma}\nabla u_2(\,\cdot\,,k) = 0,$$

where

$$\sigma(z) = (1-\mu(z))/(1+\mu(z))$$

is the conductivity corresponding to μ. Note that the conductivity $1/\sigma(z) = (1+\mu(z))/(1-\mu(z))$ is the conductivity corresponding to $-\mu$.

Generally, the near field measurements, that is, the Dirichlet-to-Neumann map Λ_σ on $\partial\Omega$, determines the scattering measurements, in particular the scattered fields outside Ω; see [Nachman 1988]. In our setting, this means that we can use Lemma 5.1 and argue, e.g., as in the proof of Proposition 6.1 in [Astala and Päivärinta 2006], that Λ_σ determines uniquely the solutions $f_{\pm\mu}(z_0, k)$ and $\tau_{\pm\mu}(k)$ for $z_0 \in \mathbb{C} \setminus \overline{\mathbb{D}}$ and $k \in \mathbb{C}$. We note that a constructive method based on integral equations on $\partial\mathbb{D}$ to determine $f_{\pm\mu}(z_0, k)$ from Λ_σ is presented in [Astala et al. 2011].

As $u_j(z, \cdot)$, $j = 1, 2$, are bounded and nonvanishing functions which satisfy (4-67), we have $\bar\partial u_j(z, \cdot) \in L^\infty_{\mathrm{loc}}(\mathbb{C})$. This implies that

$$\partial u_j(z, \cdot) \in \mathrm{BMO}_{\mathrm{loc}}(\mathbb{C}) \subset L^p_{\mathrm{loc}}(\mathbb{C}) \quad \text{for all } p < \infty$$

(see, e.g., [Astala et al. 2009, Theorem 4.6.5]), and hence $u_j(z, \cdot) \in W^{1,p}_{\mathrm{loc}}(\mathbb{C})$.

Let us now consider the isotropic conductivities σ and $\tilde\sigma$ in $\Omega = \mathbb{D}$ which are equal to 1 near $\partial\mathbb{D}$ and satisfy (1-23). Assume that $\Lambda_\sigma = \Lambda_{\tilde\sigma}$. Then, by the above considerations, $\tau_{\pm\mu}(k) = \tau_{\pm\tilde\mu}(k)$ for $k \in \mathbb{C}$.

Let

$$\mu = (1-\sigma)/(1+\sigma) \quad \text{and} \quad \tilde\mu = (1-\tilde\sigma)/(1+\tilde\sigma)$$

be the Beltrami coefficients corresponding to σ and $\tilde\sigma$.

By applying Lemma 3.3 with $k = 0$, we see that $f_\mu(z, 0) = 1$ for all $z \in \mathbb{C}$ and hence $u_1(z, 0) = 1$. By Lemma 3.2, the map $z \mapsto f_\mu(z, k)$ is continuous. Thus

$$u_1 \in \mathcal{X}^p, \quad 1 < p < \infty,$$

where $\mathcal{X}^p$ is the space of functions $v(z, k)$, $(k, z) \in \mathbb{C}^2$ for which $v(z, \cdot) \in W^{1,p}_{\mathrm{loc}}(\mathbb{C})$ and $v(z, \cdot)$ are bounded for all $z \in \mathbb{C}$ and the function $v(\cdot, k)$ is continuous for all $k \in \mathbb{C}$. These properties are crucial in the following lemma, which is a reformulation of the properties of the functions $u_1(z, k)$, with $z, k \in \mathbb{C}$, proven in [Astala and Päivärinta 2006] for L^∞-conductivities.

Lemma 4.11. (i) *The functions $u_1(z, k)$ with $k \neq 0$ have the z-asymptotics*

$$u_1(z, k) = \exp(ikz + v(z; k)), \tag{4-71}$$

where $C(k) > 0$ is such that $|v(z, k)| \le C(k)$ for all $z \in \mathbb{C}$.

(ii) *The functions $u_1(z, k)$ have the k-asymptotics*

$$u_1(z, k) = \exp(ikz + k\varepsilon_\mu(k; z)), \quad k \neq 0, \tag{4-72}$$

where for each fixed z, we have $\varepsilon_\mu(k; z) \to 0$ as $k \to \infty$.

(iii) *Let $1 < p < \infty$. The $u_1(z, k)$ given in* (4-63) *is the unique function in $\mathcal{X}^p$ such that $u_1(z, k)$ is nonvanishing, $u_1(z, 0) = 1$ for all $z \in \mathbb{C}$, and $u_1(z, k)$ satisfies the $\bar\partial$-equation* (4-67) *with the asymptotics and* (4-71) *and* (4-72).

Proof. (i) Let us omit the (z,k)-variables in some expressions and define $u_1(z,k) = u_1$, $f_\mu(z,k) = f_\mu$, etc. By the definition of u_1,

$$u_1 = \tfrac{1}{2}\big(f_\mu + f_{-\mu} + \overline{f_\mu} - \overline{f_{-\mu}}\big) = f_\mu \left(1 + \frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}}\right)^{-1} \left(1 + \frac{\overline{f_\mu} - \overline{f_{-\mu}}}{f_\mu + f_{-\mu}}\right), \tag{4-73}$$

where each factor is nonvanishing by Proposition 4.10. Thus (4-60) yields (4-71).

(ii) Let

$$F_t(z,k) = e^{-it/2}\Big(f_\mu(z,k)\cos\frac{t}{2} + i f_{-\mu}(z,k)\sin\frac{t}{2}\Big), \quad t \in \mathbb{R}.$$

Then

$$\begin{aligned} \bar\partial_z F_t(z,k) &= \mu(z) e^{-it}\,\overline{\partial_z F_t(z,k)} \quad \text{for } z \in \mathbb{C},\\ F_t(z,k) &= e^{ikz}(1 + O_k(z^{-1})) \qquad \text{as } z \to \infty. \end{aligned}$$

Thus $F_t(z,k) = \exp(k\varphi_\lambda(z,k))$, where $\lambda = e^{-it}$ and $\varphi_\lambda(z,k)$ solves (4-43). Note that $f_\mu(z,k) = \exp(k\varphi_{\lambda_0}(z,k))$, where $\lambda_0 = 1$. Then

$$\frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}} + e^{it} = \frac{2e^{it}\,F_t}{f_\mu + f_{-\mu}} = \frac{\exp(k\varphi_\lambda(z,k))}{\exp(k\varphi_{\lambda_0}(z,k))}\,\frac{2e^{it}}{1 + M_{-\mu}(z,k)/M_\mu(z,k)}. \tag{4-74}$$

By Theorem 4.9, we have, for $z \in \mathbb{C}$ and $k \in \mathbb{C} \setminus \{0\}$, that

$$e^{-|k|\varepsilon_1(k)} \le |M_{\pm\mu}(z,k)| \le e^{|k|\varepsilon_1(k)}, \tag{4-75}$$

and

$$e^{-|k|\varepsilon_2(k)} \le \inf_{|\lambda|=1}\left|\frac{\exp(k\varphi_\lambda(z,k))}{\exp(k\varphi_{\lambda_0}(z,k))}\right| \le \sup_{|\lambda|=1}\left|\frac{\exp(k\varphi_\lambda(z,k))}{\exp(k\varphi_{\lambda_0}(z,k))}\right| \le e^{|k|\varepsilon_2(k)}, \tag{4-76}$$

where $\varepsilon_j(k) \to 0$ as $k \to \infty$. Since $\operatorname{Re}(M_{-\mu}/M_\mu) > 0$, estimates (4-74) and (4-75) yield for $z \in \mathbb{C}$, $k \neq 0$, that

$$\inf_{t\in\mathbb{R}}\left|\frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}} + e^{it}\right| \ge e^{-|k|\varepsilon(k)} \quad \text{and} \quad \frac{|f_\mu - f_{-\mu}|}{|f_\mu + f_{-\mu}|} \le 1 - e^{-|k|\varepsilon(k)}.$$

This and (4-73) yield the k-asymptotics (4-72).

(iii) As observed above, the function $u_1(z,k)$ given in (4-63) satisfies the conditions stated in (iii).

Next, let $u_1(z,k)$ and $\tilde u_1(z,k)$ be two functions which satisfy the assumptions of the claim. Let us consider the logarithms

$$\delta_1(z,k) = \log u_1(z,k), \quad \tilde\delta_1(z,k) = \log \tilde u_1(z,k), \quad k, z \in \mathbb{C}.$$

As $u_1(z,\cdot) \in W^{1,p}_{\text{loc}}(\mathbb{C})$ for some $p < \infty$ and $u_1(z,\cdot)$ is a bounded and nonvanishing function, we see that $\delta_1(z,\cdot) \in W^{1,p}_{\text{loc}}(\mathbb{C})$. As $u_1(z,0) = 1$, we have

$$\delta_1(z,0) = 0 \quad \text{for } z \in \mathbb{C}. \tag{4-77}$$

Moreover, $z \mapsto \delta_1(z,k)$ is continuous for any k. Let $k \neq 0$ be fixed. Then by (4-71),

$$\delta_1(z,k) = ikz + v(z,k), \quad z \in \mathbb{C}, \tag{4-78}$$

where $v(\cdot,k)$ is bounded and we see using elementary degree theory [O'Regan et al. 2006, Corollary 1.2.10] that the map $H_k : \mathbb{C} \to \mathbb{C}$, $H_k(z) = \delta_1(z,k)$, is surjective.

The function $\tilde{\delta}_1(z,k)$ has the same above properties as $\delta_1(z,k)$. Next we want to show that $\delta_1(z,k) = \tilde{\delta}_1(z,k)$ for all $z \in \mathbb{C}$ and $k \neq 0$. As the map $H_k : z \mapsto \delta_1(z,k)$ is surjective for all $k \neq 0$, this follows if we show that

$$w \neq z \text{ and } k \neq 0 \quad \Rightarrow \quad \delta_1(w,k) \neq \tilde{\delta}_1(z,k). \tag{4-79}$$

To this end, let $z, w \in \mathbb{C}$, $z \neq w$. Functions u_1 and $\tilde{u}_1$ satisfy the same equation (4-67) with the coefficient $\tau(k) = \tau_\mu(k)$. Subtracting these equations from each other, we see that the difference $g(k; w, z) = \delta_1(w,k) - \tilde{\delta}_1(z,k)$ satisfies

$$\begin{aligned} \bar{\partial}_k g(k; w, z) &= \gamma(k; w, z)\, g(k; w, z), \quad k \in \mathbb{C}, \\ \gamma(k; w, z) &= -i\tau(k) \exp\big(i \operatorname{Im} \delta_1(k; w, z)\big) E\big(i \operatorname{Im} g(k; w, z)\big), \end{aligned} \tag{4-80}$$

where

$$E(t) = (e^{-t} - 1)/t.$$

Here, $\gamma(\cdot; w, z)$ is a locally bounded function. As $w \neq z$, the principle of the argument for pseudoanalytic functions (see [Astala and Päivärinta 2006, Proposition 3.3]), (4-80), the boundedness of γ, and the asymptotics

$$g(k; w, z) = ik(w - z) + k\varepsilon(k, w, z),$$

where $\varepsilon(k, w, z) \to 0$ as $k \to \infty$, imply that $k \mapsto g(k; w, z)$ vanishes for one and only one value of $k \in \mathbb{C}$. Thus by (4-77), $g(k; w, z) = 0$ implies that $k = 0$, and hence (4-79) holds. Thus $\delta_1(z,k) = \tilde{\delta}_1(z,k)$ and $u_1(z,k) = \tilde{u}_1(z,k)$ for all $z \in \mathbb{C}$ and $k \neq 0$. □

Remark 4.12. Note that $\tau_{\pm\mu}(k)$ is determined by Λ_σ. Thus Lemma 4.11 means that $u_1(z,k)$ can be constructed as a unique complex curve $z \mapsto u_1(z,\cdot)$, $z \in \mathbb{C}$, in the space of the solutions of the $\bar{\partial}$-equation (4-67) which has the properties stated in (iii).

When $u_j(z,k)$ and $\tilde{u}_j(z,k)$, $j = 1, 2$, are functions corresponding to μ and $\tilde{\mu}$, the above shows that $u_1(z,k) = \tilde{u}_1(z,k)$. Using $\tau_{-\mu}$ instead of τ_μ and (4-64), we see by Lemma 4.11 that $u_2(z,k) = \tilde{u}_2(z,k)$ for all $z \in \mathbb{C}$ and $k \neq 0$.

Thus $f_{\pm\mu}(z,k) = f_{\pm\tilde{\mu}}(z,k)$ for all $z \in \mathbb{C}$ and $k \neq 0$. By [Astala et al. 2009, Theorem 20.4.12], the Jacobians of $f_{\pm\mu} \in W^{1,Q}_{\text{loc}}(\mathbb{C})$ are nonvanishing almost everywhere. Thus we see using the Beltrami equation (4-59) and the fact that $f_{\pm\mu}(z,k) = f_{\pm\tilde{\mu}}(z,k)$ for all $z \in \mathbb{C}$ and $k \neq 0$ that $\mu = \tilde{\mu}$ almost everywhere. Hence $\sigma = \tilde{\sigma}$ a.e. This proves the claim of Theorem 1.9. □

5. Reduction of the inverse problem for an anisotropic conductivity to the isotropic case

In this section, we assume that the weight function $\mathcal{A}$ satisfies the almost linear growth condition (1-25). Let $\sigma = \sigma^{jk} \in \Sigma_{\mathcal{A}}(\mathbb{C})$ be a conductivity matrix such that $\sigma(z) = 1$ for z in $\mathbb{C} \setminus \Omega$ and in some neighborhood of $\partial\Omega$.

Let $z_0 \in \partial\Omega$, and define

$$\mathcal{H}_\sigma(z) = \int_{\eta_z} (\Lambda_\sigma(u|_{\partial\Omega}))(z')\, ds(z'), \tag{5-1}$$

where η_z is the path (oriented in the positive direction) from z_0 to z along $\partial\Omega$. This map is called the *σ-Hilbert transform*, and it can be considered a bounded map

$$\mathcal{H}_\sigma : H^{1/2}(\partial\Omega) \to H^{1/2}(\partial\Omega)/\mathbb{C}.$$

As shown in beginning of Section 2C, there exists a homeomorphism $F : \mathbb{C} \to \mathbb{C}$ such that $F(\Omega) = \tilde{\Omega}$, $\tilde{\sigma} = F_*\sigma$ is isotropic (i.e., a scalar function times the identity matrix), F and F^{-1} are $W^{1,P}$-smooth, and $F(z) = z + O(1/z)$. Moreover, F satisfies conditions $\mathcal{N}$ and $\mathcal{N}^{-1}$. Also, as $\sigma = 1$ near the boundary, we have that F and F^{-1} are C^∞-smooth near the boundary.

By the definition of $\tilde{\sigma} = F_*\sigma$, we see that

$$\det(\tilde{\sigma}(y)) = \det(\sigma(F^{-1}(y))) \tag{5-2}$$

for $y \in \widetilde{\Omega}$. Thus under the assumptions of Theorem 1.11, where $\det(\sigma), \det(\sigma)^{-1} \in L^\infty(\Omega)$, we see that the isotropic conductivity $\tilde{\sigma}$ satisfies $\tilde{\sigma}, \tilde{\sigma}^{-1} \in L^\infty(\Omega)$.

Let us next consider the case when the assumptions of Theorem 1.8 are valid and we have $\mathcal{A}(t) = pt - p$, with $p > 1$. Then, as F satisfies the condition $\mathcal{N}$, the area formula gives

$$\begin{aligned} I_1 &= \int_{\widetilde{\Omega}} \exp\biggl(\exp\biggl(q\biggl(\tilde{\sigma}(y) + \frac{1}{\tilde{\sigma}(y)}\biggr)\biggr)\biggr)\, dm(y) \\ &= \int_{\Omega} \exp\biggl(\exp\biggl(q\biggl(\det(\sigma(x))^{1/2} + \frac{1}{\det(\sigma(x))^{1/2}}\biggr)\biggr)\biggr) J_F(x)\, dm(x). \end{aligned} \tag{5-3}$$

In the case when $\mathcal{A}(t) = pt - p$, with $p > 1$, [Astala et al. 2010, Theorem 1.1] implies that

$$J_F \log^\beta(e + J_F) \in L^1(\Omega)$$

for $0 < \beta < p$. Then, Young's inequality (A-7) with the admissible pair (4-50) implies that

$$\begin{aligned} &\int_\Omega \exp\biggl(\exp\biggl(q\biggl(\det(\sigma(x))^{1/2} + \frac{1}{\det(\sigma(x))^{1/2}}\biggr)\biggr)\biggr) J_F(x)\, dm(x) \\ &\quad \le \biggl(\int_\Omega \exp\biggl(\exp\biggl(\exp\biggl(q\biggl(\det(\sigma)^{1/2} + \frac{1}{\det(\sigma)^{1/2}}\biggr)\biggr)\biggr)\biggr)\, dm\biggr)\biggl(\int_\Omega (1 + J_F)\log(1 + J_F)\, dm\biggr), \end{aligned} \tag{5-4}$$

and if conductivity σ satisfies (1-21), we see that I_1 is finite for some $q > 0$.

Thus under assumptions of Theorem 1.8, we see that I_1 is finite for the isotropic conductivity $\tilde{\sigma}$.

Let $\rho = F|_{\partial\Omega}$. It follows from Lemma 2.4 and (2-28) that $\rho_*\Lambda_\sigma = \Lambda_{\tilde{\sigma}}$. Then,

$$\mathcal{H}_{\tilde{\sigma}} h = \mathcal{H}_\sigma(h \circ \rho^{-1})$$

for all $h \in H^{1/2}(\partial\tilde{\Omega})$.

Next we seek a function $G_\Omega(z, k)$, with $z \in \mathbb{C} \setminus \Omega$, $k \in \mathbb{C}$, that satisfies

$$\bar{\partial}_z G_\Omega(z, k) = 0 \quad \text{for } z \in \mathbb{C} \setminus \overline{\Omega}, \tag{5-5}$$

$$G_\Omega(z, k) = e^{ikz}(1 + \mathcal{O}_k(z^{-1})) \quad \text{as } z \to \infty, \tag{5-6}$$

$$\operatorname{Im} G_\Omega(\,\cdot\,, k)|_{\partial\Omega} = \mathcal{H}_\sigma(\operatorname{Re} G_\Omega(\,\cdot\,, k)|_{\partial\Omega}). \tag{5-7}$$

To study it, we consider a similar function $G_{\widetilde{\Omega}}(\,\cdot\,,k):\mathbb{C}\setminus\widetilde{\Omega}\to\mathbb{C}$ corresponding to the scalar conductivity $\tilde{\sigma}$, which satisfies in the domain $\mathbb{C}\setminus\overline{\widetilde{\Omega}}$ the equations (5-5)–(5-6) and the boundary condition

$$\operatorname{Im} G_{\widetilde{\Omega}}(\,\cdot\,,k)=\mathcal{H}_{\tilde{\sigma}}(\operatorname{Re} G_{\widetilde{\Omega}}(\,\cdot\,,k))\qquad \text{on } z\in\partial\widetilde{\Omega}.$$

Below, let $\tilde{\mu}=(1-\tilde{\sigma})/(1+\tilde{\sigma})$ be the Beltrami coefficient corresponding to the conductivity $\tilde{\sigma}$.

Lemma 5.1. *Assume that $\sigma\in\Sigma_{\mathcal{A}}(\Omega)$ is 1 near $\partial\Omega$. Then for all $k\in\mathbb{C}$,*

(i) *For $k\in\mathbb{C}$ and $z\in\mathbb{C}\setminus\overline{\widetilde{\Omega}}$, we have $G_{\widetilde{\Omega}}(z,k)=W(z,k)$, where $W(\,\cdot\,,k)\in W^{1,P}_{\mathrm{loc}}(\mathbb{C})$ is the unique solution of*

$$\bar{\partial}_z W(z,k)=\tilde{\mu}(z)\overline{\partial_z W(z,k)}\qquad \textit{for } z\in\mathbb{C}, \tag{5-8}$$

$$W(z,k)=e^{ikz}(1+\mathcal{O}_k(z^{-1}))\quad \textit{as } z\to\infty. \tag{5-9}$$

(ii) *The equations* (5-5)–(5-7) *have a unique solution $G_{\Omega}(\,\cdot\,,k)\in C^{\infty}(\mathbb{C}\setminus\Omega)$ and $G_{\Omega}(z,k)=G_{\widetilde{\Omega}}(F(z),k)$ for $z\in\mathbb{C}\setminus\Omega$.*

Proof. The definition of the Hilbert transform $\mathcal{H}_{\tilde{\sigma}}$ implies that any solution $G_{\widetilde{\Omega}}(z,k)$ of (5-5)–(5-7) can be extended to a solution $W(z,k)$ of (5-8). On other hand, the restriction of the solution $W(z,k)$ of (5-8)–(5-9) satisfies (5-5)–(5-7). The equations (5-8)–(5-9) have a unique solution by Theorem 3.1. As the solution $W(\,\cdot\,,k)$ is analytic in $\mathbb{C}\setminus\operatorname{supp}(\tilde{\sigma})$, the claim (i) follows.

The claim (ii) follows immediately as $F:\mathbb{C}\setminus\overline{\Omega}\to\mathbb{C}\setminus\overline{\widetilde{\Omega}}$ is conformal, $F(z)=z+\mathcal{O}(1/z)$, and $\mathcal{H}_{\tilde{\sigma}}h=\mathcal{H}_{\sigma}(h\circ\rho)$ for all $h\in H^{1/2}(\partial\tilde{\Omega})$. □

Lemma 5.2. *Assume that Ω is given and that $\sigma\in\Sigma_{\mathcal{A}}(\Omega)$ is 1 near $\partial\Omega$. Then the Dirichlet-to-Neumann form L_{σ} determines the values of the restriction $F|_{\mathbb{C}\setminus\Omega}$, the boundary $\partial\widetilde{\Omega}$, and the Dirichlet-to-Neumann map $\Lambda_{\tilde{\sigma}}$ of the isotropic conductivity $\tilde{\sigma}=F_*\sigma$ on $\widetilde{\Omega}$.*

Proof. When $\sigma=1$ near $\partial\Omega$, the Dirichlet-to-Neumann form L_{σ} determines the Dirichlet-to-Neumann map Λ_{σ}. By Lemma 3.4, we have $W(z,k)=\exp(ik\varphi(z,k))$, where by Theorem 4.9,

$$\lim_{k\to\infty}\sup_{z\in\mathbb{C}}|\varphi(z,k)-z|=0. \tag{5-10}$$

For $k\neq 0$, we choose the branch of the logarithm of $G(z,k)=W(F(z),k)$ so that it is a continuous function of $z\in\mathbb{C}\setminus\Omega$ and

$$\lim_{z\to\infty}(\log G(z,k)-ikz)=0.$$

Then,

$$\lim_{k\to\infty}(ik)^{-1}\log G(z,k)=\lim_{k\to\infty}\varphi(F(z),k)=F(z). \tag{5-11}$$

By Lemma 5.1, $G(z,k)$ can be constructed for any $z\in\mathbb{C}\setminus\Omega$ by solving the equations (5-5)–(5-9). Thus the restriction of F to $\mathbb{C}\setminus\Omega$ is determined by the values of the limit (5-11). As $\widetilde{\Omega}=\mathbb{C}\setminus F(\mathbb{C}\setminus\Omega)$ and $\Lambda_{\tilde{\sigma}}=(F|_{\partial\Omega})_*\Lambda_{\sigma}$, this proves the claim. □

Above we saw that if the assumptions of Theorem 1.8 for σ are satisfied then for the isotropic conductivity $\tilde{\sigma}=F_*\sigma$, we have $\tilde{\sigma},\tilde{\sigma}^{-1}\in L^{\infty}(\widetilde{\Omega})$. Also, under the assumptions of Theorem 1.8 for σ, the integral I_1 in (5-3) is finite for some $q>0$. Thus Theorems 1.8 and 1.11 follow by Theorem 1.9 and Lemma 5.2.

Appendix: Orlicz spaces

For the proofs of the facts discussed in this appendix, we refer to [Adams 1975; Krasnosel'skiĭ and Rutickiĭ 1961].

Let $F, G : [0,\infty) \to [0,\infty)$ be bijective convex functions. The pair (F, G) is called a Young complementary pair if

$$F'(t) = f(t), \quad G'(t) = g(t), \quad g = f^{-1}.$$

In the following, we will consider also extensions of these functions defined by $F, G : \mathbb{C} \to [0,\infty)$ by setting $F(t) = F(|t|)$ and $G(t) = G(|t|)$. By [Krasnosel'skiĭ and Rutickiĭ 1961, Section I.7.4], there are examples of such pairs for which

$$F(t) = \frac{1}{p} t^p \log^a t, \quad G(t) = \frac{1}{q} t^q \log^{-a} t,$$

where $p, q \in (1,\infty)$, $p^{-1} + q^{-1} = 1$ and $a \in \mathbb{R}$. We define that $u : D \to \mathbb{C}$, where $D \subset \mathbb{R}^2$, is in the *Orlicz class* $K_F(D)$ if

$$\int_{\mathbb{D}} F(|u(x)|)\, dm(x) < \infty. \tag{A-1}$$

The Orlicz space $X_F(D)$ is the smallest vector space containing the set $K_F(D)$. For a Young complementary pair (F, G), one can define for $u \in X_F(D)$ the norm

$$\|u\|_F = \sup\left\{ \int_D |u(x)v(x)|\, dm(x) \;\middle|\; \int_D G(u(x))\, dm(x) \le 1 \right\}. \tag{A-2}$$

There is also a *Luxenburg norm*

$$\|u\|_{(F)} = \inf\left\{ t > 0 \;\middle|\; \int_D F\left(\frac{u(x)}{t}\right) dm(x) \le 1 \right\}, \tag{A-3}$$

which is equivalent to the norm $\|u\|_F$, and one always has

$$\|u\|_{(F)} \le \|u\|_F \le 2\|u\|_{(F)}. \tag{A-4}$$

By [Adams 1975, Theorem 8.10], $L_X(D)$ is a Banach space with respect to the norm $\|u\|_{(F)}$. Moreover, it holds that (see [Krasnosel'skiĭ and Rutickiĭ 1961, Theorems II.9.5 and II.10.5])

$$\|u\|_{(F)} \le 1 \quad \Rightarrow \quad \int_D F(u(x))\, dm(x) \le \|u\|_F, \tag{A-5}$$

$$\|u\|_{(F)} \ge 1 \quad \Rightarrow \quad \int_D F(u(x))\, dm(x) \ge \|u\|_{(F)}. \tag{A-6}$$

We also recall Young's inequality [Krasnosel'skiĭ and Rutickiĭ 1961, Theorem II.9.3], $uv \le F(u) + G(v)$ for $u, v \ge 0$, which implies

$$\left| \int_D u(x)v(x)\, dm(x) \right| \le \|u\|_F\, \|u\|_G. \tag{A-7}$$

The set $K_F(D)$ is a vector space when F satisfies the Δ_2-condition, that is, there is $k > 1$ such that $F(2t) \le kF(t)$ for all $t \in \mathbb{R}_+$; see [Adams 1975, Lemma 8.8]. In this case, $X_F(D) = K_F(D)$.

We will use functions

$$M_{p,q}(t) = |t|^p (\log(1 + |t|))^q, \quad 1 \le p < \infty,\ q \in \mathbb{R},$$

and use for $F(t) = M_{p,q}(t)$ the notations $X_F(D) = X^{p,q}(D)$ and $\|u\|_F = \|u\|_{X^{p,q}(D)}$. For $p = 2$, we define

$$M_{2,q}(t) = M_q(t), \quad X^{2,q}(D) = X^q(D).$$

Note that if D is bounded, $1 < p < \infty$ and $0 < \varepsilon < p - 1$, then

$$L^{p+\varepsilon}(D) \subset X^{p,q}(D) \subset L^{p-\varepsilon}(D).$$

Finally, we note that the dual space of $X^q(D)$ is $X^{-q}(D)$ and

$$\left| \int_D u(x)v(x)\, dm(x) \right| \le \|u\|_{X^q(D)} \|v\|_{X^{-q}(D)}. \tag{A-8}$$

Acknowledgements

The authors express their gratitude to the Newton Institute and the Academy of Finland. Astala was partly supported by the Finnish Centre of Excellence in Analysis and Dynamics and Lassas and Päivärinta were partly supported by the Finnish Centre of Excellence in Inverse Problems Research. Also, Päivärinta was partly supported by the European Research Council.

References

[Adams 1975] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics **65**, Academic Press, New York, 1975. MR 56 #9247 Zbl 0314.46030

[Alessandrini 1988] G. Alessandrini, "Stable determination of conductivity by boundary measurements", *Appl. Anal.* **27**:1-3 (1988), 153–172. MR 89f:35195 Zbl 0616.35082

[Alessandrini 1990] G. Alessandrini, "Singular solutions of elliptic equations and the determination of conductivity by boundary measurements", *J. Differential Equations* **84**:2 (1990), 252–272. MR 91e:35210 Zbl 0778.35109

[Alu and Engheta 2005] A. Alu and N. Engheta, "Achieving transparency with plasmonic and metamaterial coatings", *Phys. Rev. E* **72**:1 (2005), Article ID #016623. Erratum in **73**:1 (2006), Article ID #019906.

[Alu and Engheta 2009] A. Alu and N. Engheta, "Cloaking a sensor", *Phys. Rev. Lett.* **102** (2009), Article ID #233901.

[Ammari et al. 2013] H. Ammari, H. Kang, H. Lee, and M. Lim, "Enhancement of near cloaking using generalized polarization tensors vanishing structures, I: The conductivity problem", *Comm. Math. Phys.* **317**:1 (2013), 253–266. MR 3010374 Zbl 1303.35108

[Astala 1994] K. Astala, "Area distortion of quasiconformal mappings", *Acta Math.* **173**:1 (1994), 37–60. MR 95m:30028b Zbl 0815.30015

[Astala and Päivärinta 2006] K. Astala and L. Päivärinta, "Calderón's inverse conductivity problem in the plane", *Ann. of Math.* (2) **163**:1 (2006), 265–299. MR 2007b:30019 Zbl 1111.35004

[Astala et al. 2005] K. Astala, L. Päivärinta, and M. Lassas, "Calderón's inverse problem for anisotropic conductivity in the plane", *Comm. Partial Differential Equations* **30**:1-3 (2005), 207–224. MR 2005k:35421 Zbl 1129.35483

[Astala et al. 2009] K. Astala, T. Iwaniec, and G. Martin, *Elliptic partial differential equations and quasiconformal mappings in the plane*, Princeton Mathematical Series **48**, Princeton University Press, 2009. MR 2010j:30040 Zbl 1182.30001

[Astala et al. 2010] K. Astala, J. T. Gill, S. Rohde, and E. Saksman, "Optimal regularity for planar mappings of finite distortion", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**:1 (2010), 1–19. MR 2011k:30025 Zbl 1191.30007

[Astala et al. 2011] K. Astala, J. L. Mueller, L. Päivärinta, A. Perämäki, and S. Siltanen, "Direct electrical impedance tomography for nonsmooth conductivities", *Inverse Probl. Imaging* **5**:3 (2011), 531–549. MR 2012g:65238 Zbl 1237.78014

[Attouch et al. 2006] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*, MPS/SIAM Series on Optimization **6**, Society for Industrial and Applied Mathematics, Philadelphia, 2006. MR 2006j:49001 Zbl 1095.49001

[Ball 1982] J. M. Ball, "Discontinuous equilibrium solutions and cavitation in nonlinear elasticity", *Philos. Trans. Roy. Soc. London Ser. A* **306**:1496 (1982), 557–611. MR 84i:73041 Zbl 0513.73020

[Barceló et al. 2001] J. A. Barceló, T. Barceló, and A. Ruiz, "Stability of the inverse conductivity problem in the plane for less regular conductivities", *J. Differential Equations* **173**:2 (2001), 231–270. MR 2002h:35325 Zbl 0986.35126

[Barceló et al. 2007] T. Barceló, D. Faraco, and A. Ruiz, "Stability of Calderón inverse conductivity problem in the plane", *J. Math. Pures Appl.* (9) **88**:6 (2007), 522–556. MR 2009a:35255 Zbl 1133.35104

[Borcea 1999] L. Borcea, "Asymptotic analysis of quasi-static transport in high contrast conductive media", *SIAM J. Appl. Math.* **59**:2 (1999), 597–635. MR 2001a:78006 Zbl 0927.35028

[Borcea 2002] L. Borcea, "Electrical impedance tomography", *Inverse Probl.* **18**:6 (2002), R99–R136. MR 1955896 Zbl 1031.35147

[Borcea et al. 1996] L. Borcea, J. G. Berryman, and G. C. Papanicolaou, "High-contrast impedance tomography", *Inverse Probl.* **12**:6 (1996), 835–858. MR 1421651 Zbl 0862.35137

[Brown and Uhlmann 1997] R. M. Brown and G. A. Uhlmann, "Uniqueness in the inverse conductivity problem for non-smooth conductivities in two dimensions", *Comm. Partial Differential Equations* **22**:5-6 (1997), 1009–1027. MR 98f:35155 Zbl 0884.35167

[Calderón 1980] A.-P. Calderón, "On an inverse boundary value problem", pp. 65–73 in *Seminar on Numerical Analysis and its Applications to Continuum Physics* (Rio de Janeiro, 1980), Coleção Atas **12**, Sociedade Brasileira de Matemática, Rio de Janeiro, 1980. Reprinted in *Comput. Appl Math* **25**:2–3 (2006), 133–138. MR 81k:35160 Zbl 1182.35230

[Capolino 2009] F. Capolino, *Applications of metamaterials*, CRC Press, Boca Raton, FL, 2009.

[Chen and Chan 2007a] H. Chen and C. T. Chan, "Acoustic cloaking in three dimensions using acoustic metamaterials", *Appl. Phys. Lett.* **91**:18 (2007), Article ID #183518.

[Chen and Chan 2007b] H. Chen and C. T. Chan, "Transformation media that rotate electromagnetic fields", *Appl. Phys. Lett.* **90**:24 (2007), Article ID #241105.

[Colton and Kress 2013] D. Colton and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, 3rd ed., Applied Mathematical Sciences **93**, Springer, New York, 2013. MR 2986407 Zbl 1266.35121

[Dos Santos Ferreira et al. 2009] D. Dos Santos Ferreira, C. E. Kenig, M. Salo, and G. A. Uhlmann, "Limiting Carleman weights and anisotropic inverse problems", *Invent. Math.* **178**:1 (2009), 119–171. MR 2010h:58033 Zbl 1181.35327

[Ergin et al. 2010] T. Ergin et al., "Three-dimensional invisibility cloak at optical wavelengths", *Science* **328**:5976 (2010), 337–339.

[Gömöry et al. 2012] F. Gömöry et al., "Experimental realization of a magnetic cloak", *Science* **335**:6075 (2012), 1466–1468.

[Greenleaf et al. 2003a] A. Greenleaf, M. Lassas, and G. A. Uhlmann, "Anisotropic conductivities that cannot be detected by EIT", *Physiolog. Meas.* **24**:2 (2003), 413–420.

[Greenleaf et al. 2003b] A. Greenleaf, M. Lassas, and G. A. Uhlmann, "The Calderón problem for conormal potentials, I: Global uniqueness and reconstruction", *Comm. Pure Appl. Math.* **56**:3 (2003), 328–352. MR 2003j:35324 Zbl 1061.35165

[Greenleaf et al. 2003c] A. Greenleaf, M. Lassas, and G. A. Uhlmann, "On nonuniqueness for Calderón's inverse problem", *Math. Res. Lett.* **10**:5-6 (2003), 685–693. MR 2005f:35316 Zbl 1054.35127

[Greenleaf et al. 2007a] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Effectiveness and improvement of cylindrical cloaking with the SHS lining", *Optics Express* **15** (2007), 12717–12734.

[Greenleaf et al. 2007b] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Electromagnetic wormholes and virtual magnetic monopoles from metamaterials", *Phys. Rev. Lett.* **99**:18 (2007), Article ID #183901.

[Greenleaf et al. 2007c] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Full-wave invisibility of active devices at all frequencies", *Comm. Math. Phys.* **275**:3 (2007), 749–789. MR 2008g:78016 Zbl 1151.78006

[Greenleaf et al. 2008] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Approximate quantum cloaking and almost-trapped states", *Phys. Rev. Lett.* **101**:22 (2008), Article ID #220404.

[Greenleaf et al. 2009a] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Cloaking devices, electromagnetic wormholes, and transformation optics", *SIAM Rev.* **51**:1 (2009), 3–33. MR 2010b:35484 Zbl 1158.78004

[Greenleaf et al. 2009b] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Invisibility and inverse problems", *Bull. Amer. Math. Soc.* (*N.S.*) **46**:1 (2009), 55–97. MR 2010d:35399 Zbl 1159.35074

[Greenleaf et al. 2011a] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Approximate quantum and acoustic cloaking", *J. Spectr. Theory* **1**:1 (2011), 27–80. MR 2820885 Zbl 1257.35156

[Greenleaf et al. 2011b] A. Greenleaf, Y. Kurylev, M. Lassas, and G. A. Uhlmann, "Cloaking a sensor via transformation optics", *Phys. Rev. E* **83**:1 (2011), Article ID #016603. MR 2012a:78003

[Greenleaf et al. 2012] A. Greenleaf, Y. Kurylev, M. Lassas, U. Leonhardt, and G. A. Uhlmann, "Schrödinger's hat: electromagnetic, acoustic and quantum amplifiers via transformation optics", *Proc. Nat. Acad. Sci.* **109** (2012), 10169–10174.

[Haberman and Tataru 2013] B. Haberman and D. Tataru, "Uniqueness in Calderón's problem with Lipschitz conductivities", *Duke Math. J.* **162**:3 (2013), 496–516. MR 3024091 Zbl 1260.35251

[Hencl et al. 2005] S. Hencl, P. Koskela, and J. Onninen, "A note on extremal mappings of finite distortion", *Math. Res. Lett.* **12**:2-3 (2005), 231–237. MR 2006a:30022 Zbl 1079.30024

[Imanuvilov et al. 2010] O. Y. Imanuvilov, G. A. Uhlmann, and M. Yamamoto, "The Calderón problem with partial data in two dimensions", *J. Amer. Math. Soc.* **23**:3 (2010), 655–691. MR 2012c:35472 Zbl 1201.35183

[Ingerman 2000] D. V. Ingerman, "Discrete and continuous Dirichlet-to-Neumann maps in the layered case", *SIAM J. Math. Anal.* **31**:6 (2000), 1214–1234. MR 2001g:39043 Zbl 0972.35183

[Isaacson et al. 2004] D. Isaacson, J. L. Mueller, J. Newell, and S. Siltanen, "Reconstructions of chest phantoms by the D-bar method for electrical impedance tomography", *IEEE Trans. Med. Im.* **23**:7 (2004), 821–828.

[Iwaniec and Martin 2001] T. Iwaniec and G. Martin, *Geometric function theory and non-linear analysis*, Clarendon/Oxford University Press, New York, 2001. MR 2003c:30001 Zbl 1045.30011

[Kauhanen et al. 2003] J. Kauhanen, P. Koskela, J. Malý, J. Onninen, and X. Zhong, "Mappings of finite distortion: sharp Orlicz-conditions", *Rev. Mat. Iberoamericana* **19**:3 (2003), 857–872. MR 2005a:30038 Zbl 1059.30017

[Kenig et al. 2007] C. E. Kenig, J. Sjöstrand, and G. A. Uhlmann, "The Calderón problem with partial data", *Ann. of Math.* (2) **165**:2 (2007), 567–591. MR 2008k:35498 Zbl 1127.35079

[Kilpeläinen et al. 2000] T. Kilpeläinen, J. Kinnunen, and O. Martio, "Sobolev spaces with zero boundary values on metric spaces", *Potential Anal.* **12**:3 (2000), 233–247. MR 2000m:46071 Zbl 0962.46021

[Knudsen et al. 2007] K. Knudsen, M. Lassas, J. L. Mueller, and S. Siltanen, "D-bar method for electrical impedance tomography with discontinuous conductivities", *SIAM J. Appl. Math.* **67**:3 (2007), 893–913. MR 2007m:35276 Zbl 1123.35091

[Knudsen et al. 2009] K. Knudsen, M. Lassas, J. L. Mueller, and S. Siltanen, "Regularized D-bar method for the inverse conductivity problem", *Inverse Probl. Imaging* **3**:4 (2009), 599–624. MR 2011d:35518 Zbl 1184.35314

[Kohn and Vogelius 1984] R. V. Kohn and M. S. Vogelius, "Identification of an unknown conductivity by means of measurements at the boundary", pp. 113–123 in *Inverse problems* (New York, 1983), edited by D. W. McLaughlin, SIAM-AMS Proceedings **14**, American Mathematical Society, Providence, RI, 1984. MR 773707 Zbl 0573.35084

[Kohn et al. 2008] R. V. Kohn, H. Shen, M. S. Vogelius, and M. I. Weinstein, "Cloaking via change of variables in electric impedance tomography", *Inverse Probl.* **24**:1 (2008), Article ID #015016. MR 2008m:78014 Zbl 1153.35406

[Kohn et al. 2010] R. V. Kohn, D. Onofrei, M. S. Vogelius, and M. I. Weinstein, "Cloaking via change of variables for the Helmholtz equation", *Comm. Pure Appl. Math.* **63**:8 (2010), 973–1016. MR 2011j:78004 Zbl 1194.35505

[Kolehmainen et al. 2005] V. Kolehmainen, M. Lassas, and P. Ola, "The inverse conductivity problem with an imperfectly known boundary", *SIAM J. Appl. Math.* **66**:2 (2005), 365–383. MR 2007f:35298 Zbl 1141.35472

[Kolehmainen et al. 2010] V. Kolehmainen, M. Lassas, and P. Ola, "Calderón's inverse problem with an imperfectly known boundary and reconstruction up to a conformal deformation", *SIAM J. Math. Anal.* **42**:3 (2010), 1371–1381. MR 2011g:35427 Zbl 1217.30020

[Kolehmainen et al. 2013] V. Kolehmainen, M. Lassas, P. Ola, and S. Siltanen, "Recovering boundary shape and conductivity in electrical impedance tomography", *Inverse Probl. Imaging* **7**:1 (2013), 217–242. MR 3031845 Zbl 1267.35254

[Krasnosel'skiĭ and Rutickiĭ 1961] M. A. Krasnosel'skiĭ and J. B. Rutickiĭ, *Convex functions and Orlicz spaces*, Noordhoff, Groningen, 1961. MR 23 #A4016 Zbl 0095.09103

[Lai et al. 2009] Y. Lai et al., "Illusion optics: the optical transformation of an object into another object", *Phys. Rev. Lett.* **102**:25 (2009), Article ID #253902.

[Landy et al. 2008] N. Landy et al., "Perfect metamaterial absorber", *Phys. Rev. Lett.* **100**:20 (2008), Article ID #207402.

[Lassas and Uhlmann 2001] M. Lassas and G. A. Uhlmann, "On determining a Riemannian manifold from the Dirichlet-to-Neumann map", *Ann. Sci. École Norm. Sup.* (4) **34**:5 (2001), 771–787. MR 2003e:58037 Zbl 0992.35120

[Lassas and Zhou 2011] M. Lassas and T. Zhou, "Two dimensional invisibility cloaking for Helmholtz equation and non-local boundary conditions", *Math. Res. Lett.* **18**:3 (2011), 473–488. MR 2012d:35062 Zbl 1241.35045

[Lassas et al. 2003] M. Lassas, M. Taylor, and G. A. Uhlmann, "The Dirichlet-to-Neumann map for complete Riemannian manifolds with boundary", *Comm. Anal. Geom.* **11**:2 (2003), 207–221. MR 2004h:58033 Zbl 1077.58012

[Lee and Uhlmann 1989] J. M. Lee and G. A. Uhlmann, "Determining anisotropic real-analytic conductivities by boundary measurements", *Comm. Pure Appl. Math.* **42**:8 (1989), 1097–1112. MR 91a:35166 Zbl 0702.35036

[Leonhardt 2006] U. Leonhardt, "Optical conformal mapping", *Science* **312**:5781 (2006), 1777–1780. MR 2237569 Zbl 1226.78001

[Liu 2013] H. Liu, "On near-cloak in acoustic scattering", *J. Differential Equations* **254**:3 (2013), 1230–1246. MR 2997368 Zbl 1257.35173

[Liu and Sun 2013] H. Liu and H. Sun, "Enhanced near-cloak by FSH lining", *J. Math. Pures Appl.* (9) **99**:1 (2013), 17–42. MR 3003281 Zbl 1259.35223

[Liu and Zhou 2011] H. Liu and T. Zhou, "On approximate electromagnetic cloaking by transformation media", *SIAM J. Appl. Math.* **71**:1 (2011), 218–241. MR 2012f:78005 Zbl 1213.35403

[Manfredi 1994] J. J. Manfredi, "Weakly monotone functions", *J. Geom. Anal.* **4**:3 (1994), 393–402. MR 95m:30032 Zbl 0805.35013

[Milton and Nicorovici 2006] G. W. Milton and N.-A. P. Nicorovici, "On the cloaking effects associated with anomalous localized resonance", *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **462**:2074 (2006), 3027–3059. MR 2008e:78018 Zbl 1149.00310

[Morrey 1938] C. B. Morrey, Jr., "On the solutions of quasi-linear elliptic partial differential equations", *Trans. Amer. Math. Soc.* **43**:1 (1938), 126–166. MR 1501936 Zbl 0018.40501

[Mueller and Siltanen 2012] J. L. Mueller and S. Siltanen, *Linear and nonlinear inverse problems with practical applications*, Computational Science and Engineering **10**, Society for Industrial and Applied Mathematics, Philadelphia, 2012. MR 2986262 Zbl 1262.65124

[Nachman 1988] A. I. Nachman, "Reconstructions from boundary measurements", *Ann. of Math.* (2) **128**:3 (1988), 531–576. MR 90i:35283 Zbl 0675.35084

[Nachman 1996] A. I. Nachman, "Global uniqueness for a two-dimensional inverse boundary value problem", *Ann. of Math.* (2) **143**:1 (1996), 71–96. MR 96k:35189 Zbl 0857.35135

[Nguyen 2012] H.-M. Nguyen, "Approximate cloaking for the Helmholtz equation via transformation optics and consequences for perfect cloaking", *Comm. Pure Appl. Math.* **65**:2 (2012), 155–186. MR 2855543 Zbl 1231.35310

[O'Regan et al. 2006] D. O'Regan, Y. J. Cho, and Y.-Q. Chen, *Topological degree theory and applications*, Series in Mathematical Analysis and Applications **10**, Chapman & Hall/CRC, Boca Raton, FL, 2006. MR 2007b:47168 Zbl 1095.47001

[Päivärinta et al. 2003] L. Päivärinta, A. Panchenko, and G. A. Uhlmann, "Complex geometrical optics solutions for Lipschitz conductivities", *Rev. Mat. Iberoamericana* **19**:1 (2003), 57–72. MR 2004f:35187 Zbl 1055.35144

[Pendry et al. 2006] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling electromagnetic fields", *Science* **312**:5781 (2006), 1780–1782. MR 2237570 Zbl 1226.78003

[Schurig et al. 2006] D. Schurig, J. Mock, B. Justice, S. Cummer, J. B. Pendry, A. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies", *Science* **314**:5801 (2006), 977–980.

[Siltanen et al. 2000] S. Siltanen, J. L. Mueller, and D. Isaacson, "An implementation of the reconstruction algorithm of A. Nachman for the 2D inverse conductivity problem", *Inverse Probl.* **16**:3 (2000), 681–699. MR 2001g:35269 Zbl 0962.35193

[Sylvester 1990] J. Sylvester, "An anisotropic inverse boundary value problem", *Comm. Pure Appl. Math.* **43**:2 (1990), 201–232. MR 90m:35202 Zbl 0709.35102

[Sylvester and Uhlmann 1987] J. Sylvester and G. A. Uhlmann, "A global uniqueness theorem for an inverse boundary value problem", *Ann. of Math.* (2) **125**:1 (1987), 153–169. MR 88b:35205 Zbl 0625.35078

[Yang et al. 2012] F. Yang, Z. Mei, T. Jin, and T. Cui, "dc electric invisibility cloak", *Phys. Rev. Lett.* **109**:5 (2012), Article ID #053902.

Received 2 Jul 2015. Accepted 3 Sep 2015.

KARI ASTALA: kari.astala@helsinki.fi
Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FI-00014 Helsinki, Finland

MATTI LASSAS: matti.lassas@helsinki.fi
Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FI-00014 Helsinki, Finland

LASSI PÄIVÄRINTA: lassi.paivarinta@ttu.ee
Department of Mathematics, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.99

A CHARACTERIZATION OF 1-RECTIFIABLE DOUBLING MEASURES WITH CONNECTED SUPPORTS

JONAS AZZAM AND MIHALIS MOURGOGLOU

Garnett, Killip, and Schul have exhibited a doubling measure μ with support equal to $\mathbb{R}^d$ that is 1-*rectifiable*, meaning there are countably many curves Γ_i of finite length for which $\mu\big(\mathbb{R}^d \setminus \bigcup \Gamma_i\big) = 0$. In this note, we characterize when a doubling measure μ with support equal to a connected metric space X has a 1-rectifiable subset of positive measure and show this set coincides up to a set of μ-measure zero with the set of $x \in X$ for which $\liminf_{r\to 0} \mu(B_X(x,r))/r > 0$.

1. Introduction

Recall that a Borel measure μ on a metric space X is *doubling* if there is $C_\mu > 0$ so that

$$\mu(B_X(x,2r)) \le C_\mu \mu(B_X(x,r)) \quad \text{for all } x \in X \text{ and } r > 0. \tag{1-1}$$

Garnett, Killip, and Schul [Garnett et al. 2010] exhibit a doubling measure μ with support equal to $\mathbb{R}^n$, $n > 1$, that is 1-rectifiable in the sense that there are countably many curves Γ_i of finite length such that $\mu\big(\mathbb{R}^n \setminus \bigcup \Gamma_i\big) = 0$. This is surprising given that such measures give zero measure to smooth or bi-Lipschitz curves in $\mathbb{R}^d$. To see this, note that, for such a curve Γ and for each $x \in \Gamma$, there are $r_x, \delta_x > 0$ so that for all $r \in (0, r_x)$ there is $B_{\mathbb{R}^d}(y_{x,r}, \delta_x r) \subseteq B_{\mathbb{R}^n}(x, r_x) \setminus \Gamma$, so by the Lebesgue differentiation theorem, $\mu(\Gamma) = 0$. If Γ is just Lipschitz and not bi-Lipschitz, however, we only know this property holds for every point in Γ outside a set of zero length. The aforementioned result shows that Lipschitz curves of finite length can in some sense be coiled up tightly enough that this zero-length set accumulates on a set of positive doubling measure.

The notion of rectifiability of a measure that we are using is not universal. In [Azzam et al. 2015], a measure μ in Euclidean space being d-rectifiable means $\mu \ll \mathscr{H}^d$ and $\operatorname{supp}\mu$ is d-rectifiable. In our setting, however, we don't require absolute continuity of our measures. To avoid ambiguity, we fix our definition below, which is the convention used in [Federer 1969, §3.2.14].

Definition 1.1. If μ is a Borel measure on a metric space X, d is an integer, and $E \subseteq X$ a Borel set, we say E is (μ, d)-*rectifiable* if $\mu\big(E \setminus \bigcup_{i=1}^\infty \Gamma_i\big) = 0$ where $\Gamma_i = f_i(E_i)$, $E_i \subseteq \mathbb{R}^d$, and $f_i : E_i \to X$ is Lipschitz. We say μ is d-*rectifiable* if $\operatorname{supp}\mu$ is (μ, d)-rectifiable.

A set $E \subseteq \mathbb{R}^n$ of positive and finite $\mathscr{H}^d$-measure is d-*rectifiable* if it is $(\mathscr{H}^d, d)$-rectifiable (see [Mattila 1995, Definition 15.3] and the few paragraphs preceding it). This is also equivalent to being covered up

The authors were supported by the ERC grant 320501 of the European Research Council (FP7/2007–2013).
MSC2010: primary 28A75; secondary 28A78.
Keywords: doubling measures, rectifiability, porosity, connected metric spaces.

to set of $\mathscr{H}^d$-measure zero by Lipschitz graphs [Mattila 1995, Lemma 15.4]. The example from [Garnett et al. 2010], however, shows that being almost covered by Lipschitz graphs versus Lipschitz images are not equivalent definitions for rectifiability of a measure.

Since this example was published, it has been an open question to classify which doubling measures on $\mathbb{R}^d$ are rectifiable. Very recently, Badger and Schul have given a complete description. First, for a general Radon measure in $\mathbb{R}^d$ and A compact with $\mu(A) > 0$, define

$$\beta_2^{(1)}(\mu, A)^2 = \inf_L \int_A \left(\frac{\operatorname{dist}(x, L)}{\operatorname{diam} A} \right)^2 \frac{d\mu(x)}{\mu(A)}$$

where the infimum is taken over all lines $L \subseteq \mathbb{R}^d$.

Theorem 1.2 [Badger and Schul 2015b, Corollary 1.12]. *If μ is a Radon measure on $\mathbb{R}^d$ such that $\liminf_{r\to 0} \beta_2^{(1)}(\mu, B_{\mathbb{R}^d}(x, r)) > 0$ for μ-almost every $x \in \mathbb{R}^d$, then μ is 1-rectifiable if and only if*

$$\sum_{\substack{x\in Q \\ \ell(Q)\le 1}} \frac{\operatorname{diam} Q}{\mu(Q)} < \infty \quad \mu\text{-a.e.} \tag{1-2}$$

where the sum is over half-open dyadic cubes Q.

It is not hard to show that, if μ is a doubling measure with $\operatorname{supp} \mu = \mathbb{R}^d$, $d \ge 2$, then there is $c > 0$ depending on the doubling constant such that $\beta_2^{(1)}(\mu, B) \ge c > 0$ for any ball $B \subseteq \mathbb{R}^d$, so the above theorem characterizes all 1-rectifiable doubling measures with support equal to all of $\mathbb{R}^d$.

In this short note, we take a different approach and provide a complete classification of 1-rectifiable doubling measures not just with support equal to $\mathbb{R}^d$ but with support equal to any topologically connected metric space. It turns out that the rectifiable part of such a measure coincides up to a set of μ-measure zero with the set of points where the lower 1-density is positive, where for $s > 0$ we define the *lower s-density* as

$$\underline{D}^s(\mu, x) := \liminf_{r\to 0} \frac{\mu(B_X(x, r))}{r^s}.$$

Theorem 1.3 (main theorem). *Let μ be a doubling measure whose support is a topologically connected metric space X, and let $E \subseteq X$ be compact. Then E is $(\mu, 1)$-rectifiable if and only if $\underline{D}^1(\mu, x) > 0$ for μ-a.e. $x \in E$.*

Note that there are no other topological or geometric restrictions on X: the support of μ may have topological dimension two (like $\mathbb{R}^2$ for example), yet if $\underline{D}^1(\mu, x) > 0$ μ-a.e., then μ is supported on a countable union of Lipschitz images of $\mathbb{R}$. Also observe that the condition $\underline{D}^1(\mu, x) > 0$ is a weaker condition than (1-2). An interesting corollary of the main theorem and Theorem 1.2 is the following.

Corollary 1.4. *If μ is a doubling measure in $\mathbb{R}^d$ with connected support such that*

$$\liminf_{r\to 0} \beta_2^{(1)}(\mu, B_{\mathbb{R}^d}(x, r)) > 0$$

and $\underline{D}^1(\mu, x) > 0$ μ-a.e., then (1-2) holds.

2. Proof of the main theorem: sufficiency

When dealing with any metric space X, we will let $B_X(x, r)$ denote the set of points *in* X of distance less than $r > 0$ from x. If $B = B_X(x, r)$ and $M > 0$, we will denote $MB = B_X(x, Mr)$. For a Borel set $A \subseteq X$, we define the (spherical) 1-Hausdorff measure as

$$\mathscr{H}^1_\delta(A) = \inf\left\{\sum_{i=1}^{\infty} 2r_i : A \subseteq \bigcup_{i=1}^{\infty} B_X(x_i, r_i),\ x_i \in A,\ r_i \in (0, \delta)\right\}$$

and $\mathscr{H}^1(A) = \inf_{\delta>0} \mathscr{H}^1_\delta(A)$.

For $A, B \subseteq X$, we set

$$\operatorname{dist}(A, B) = \inf\{|x - y| : x \in A,\ y \in B\}$$

and, for $x \in X$, $\operatorname{dist}(x, A) = \operatorname{dist}(\{x\}, A)$.

Remark 2.1. By the Kuratowski embedding theorem, if X is separable (which happens, for example, if $X = \operatorname{supp} \mu$ for a locally finite measure μ), X is isometrically embeddable into $C(X)$, where $C(X)$ is the Banach space of bounded continuous functions on X equipped with the supremum norm $|f| = \sup_{x\in X} |f(x)|$. Thus, we can assume without loss of generality that X is the subset of a complete Banach space, and we will abuse notation by calling this space $C(X)$ as well so that $X \subseteq C(X)$.

The forward direction of the main theorem is proven for general measures in Euclidean space by Badger and Schul [2015a, Lemma 2.7], who in fact prove a higher-dimensional version. Below we provide a proof that works for metric spaces in the one-dimensional case.

Proposition 2.2. *Let μ be a finite measure with $X := \operatorname{supp} \mu$ a metric space, and suppose μ is 1-rectifiable. Then $\underline{D}^1(\mu, x) > 0$ for μ-a.e. $x \in \operatorname{supp} \mu$.*

Proof. Let

$$F = \{x \in \operatorname{supp} \mu : \underline{D}^1(\mu, x) = 0\},$$

and let $\varepsilon, \delta > 0$. Since μ is rectifiable, there are Lipschitz functions $f_i : A_i \to X$, where $A_i \subseteq [0, 1]$ are compact Borel sets of positive measure and $i = 1, \ldots, N$, so that

$$\mu\left(E \setminus \bigcup_{i=1}^{N} f_i(A_i)\right) < \delta.$$

We can extend each f_i affinely on the intervals in the complement of A_i to a Lipschitz function $f_i : [0, 1] \to C(X)$. Let $d = \min_{i=1,\ldots,N} \operatorname{diam} f_i([0, 1])$ so that $r \in (0, d)$ and $x \in G := \bigcup_{i=1}^{N} f_i([0, 1])$ implies $\mathscr{H}^1(B_{C(X)}(x, r) \cap G) \geq r$ (simply because now the images of the f_i are connected).

For each $x \in F \cap G$, there is $r_x \in (0, d/5)$ so that $\mu(B_X(x, 5r_x)) < \varepsilon r_x$. By the Vitali covering theorem [Heinonen 2001, Lemma 1.2], there are countably many disjoint balls $B_i = B_X(x_i, r_i)$ with centers in F so that $\bigcup 5B_i \supseteq F$. Thus,

$$\mu(F \cap G) \leq \sum_i \mu(5B_i) \leq \varepsilon \sum_i r_i \leq \varepsilon \sum_i \mathscr{H}^1(B_{C(X)}(x_i, r_i) \cap G) \leq \varepsilon \mathscr{H}^1(G).$$

Thus,

$$\mu(F) < \delta + \varepsilon \mathscr{H}^1(G).$$

Keeping δ (and hence G) fixed and sending $\varepsilon \to 0$, we get $\mu(F) < \delta$ for all $\delta > 0$ and thus $\mu(F) = 0$. □

3. Proof of the main theorem: necessity

What remains is to prove the reverse direction of the main theorem, which we summarize in the next lemma.

Lemma 3.1. *Let μ be a doubling measure with constant $C_\mu > 0$ and support X, a topologically connected metric space. Then $\{x \in X : \underline{D}^1(\mu, x) > 0\}$ is $(\mu, 1)$-rectifiable.*

To prove Lemma 3.1, it suffices to show the following lemma.

Lemma 3.2. *Let μ be a doubling measure and support X a topologically connected complete metric space. If $E \subseteq X$ is a compact set for which $E \subseteq B_X(\xi_0, r_0/2)$ for some $\xi_0 \in X$, $r_0 > 0$, and*

$$\mu(B_X(x, r)) \geq 2r \quad \text{for all } x \in E \text{ and } r \in (0, r_0), \tag{3-1}$$

then $E = f(A)$ for some $A \subseteq \mathbb{R}$ and Lipschitz function $f : A \to X$.

Proof of Lemma 3.1 using Lemma 3.2. First, note that, if we define $\bar{\mu}(A) = \mu(A \cap X)$, then $\bar{\mu}$ is a doubling measure on $\overline{X}$, where the closure is in $C(X)$ (recall Remark 2.1). Moreover, the closure $\overline{X}$ is still topologically connected but now is a complete metric space since $C(X)$ is complete. Thus, for proving Lemma 3.1, we can assume without loss of generality that X is complete.

Let $F := \{x \in X : \underline{D}^1(\mu, x) > 0\}$. For $j, k \in \mathbb{N}$, let

$$F_{j,k} = \{x \in F : \mu(B_X(x, r)) \geq r/j \text{ for } 0 < r < k^{-1}\}.$$

Then $F = \bigcup_{j,k \in \mathbb{N}} F_{j,k}$. Furthermore, we can write $F_{j,k}$ as a countable union of sets $\{F_{j,k,\ell}\}_{\ell \in \mathbb{N}}$ with diameters less than $1/(3k)$. It suffices then to show that each one of these sets is 1-rectifiable. Fix $j, k, \ell \in \mathbb{N}$. Then the measure $j\mu$ and the set $F_{j,k,\ell}$ satisfy the conditions for Lemma 3.2 with $r_0 = k^{-1}$ except that $F_{j,k,\ell}$ is not necessarily compact. However, $\overline{F}_{j,k,\ell}$ is a closed set still satisfying these conditions, it is totally bounded since μ is doubling, and since X is complete, the Heine–Borel theorem implies $\overline{F}_{j,k,\ell}$ is compact. Thus, we can apply Lemma 3.2 to get that $\overline{F}_{j,k,\ell}$ is rectifiable. Since $F = \bigcup_{j,k,\ell} F_{j,k,\ell}$, we now have that F is also rectifiable. □

The rest of the paper is devoted to proving Lemma 3.2, so fix μ, E, ξ_0, and r_0 as in the lemma.

Proof of Lemma 3.2. We will require the notion of dyadic cubes on a metric space. This theorem was originally developed by David [1988] and Christ [1990], but the current formulation we take from Hytönen and Martikainen [2012].

Theorem 3.3. *Let X be a metric space equipped with a doubling measure μ. Let X_n be a nested sequence of maximal ρ^n-nets for X where $\rho < 1/1000$, and let $c_0 = 1/500$. For each $n \in \mathbb{Z}$, there is a collection $\mathscr{D}_n$ of "cubes", which are Borel subsets of X such that:*

(1) *For every* n, $X=\bigcup_{\Delta\in\mathscr{D}_n}\Delta$.

(2) *If* $\Delta,\Delta'\in\mathscr{D}=\bigcup\mathscr{D}_n$ *and* $\Delta\cap\Delta'\neq\varnothing$, *then* $\Delta\subseteq\Delta'$ *or* $\Delta'\subseteq\Delta$.

(3) *For* $\Delta\in\mathscr{D}$, *let* $n(\Delta)$ *be the unique integer so that* $\Delta\in\mathscr{D}_n$ *and set* $\ell(\Delta)=5\rho^{n(\Delta)}$. *Then there is* $\zeta_\Delta\in X_n$ *so that*
$$B_X(\zeta_\Delta,c_0\ell(\Delta))\subseteq\Delta\subseteq B_X(\zeta_\Delta,\ell(\Delta))$$
and
$$X_n=\{\zeta_\Delta:\Delta\in\mathscr{D}_n\}.$$

It is not necessary for there to exist a doubling measure but just that the metric space is geometrically doubling. Moreover, Hytönen and Martikainen [2012] use sequences of sets X_n slightly more general than maximal nets.

Let X_n be a nested sequence of maximal ρ^n-nets for X where $\rho<1/1000$ and $\mathscr{D}$ the resulting cubes from Theorem 3.3. By picking our net points X_n appropriately, we may assume that $E\subseteq\Delta_0\in\mathscr{D}$.

Lemma 3.4 [Azzam 2014, §3]. *Let* μ *be a* C_μ*-doubling measure and* $\mathscr{D}$ *the cubes from Theorem 3.3 for* $X=\operatorname{supp}\mu$ *with admissible constants* c_0 *and* ρ. *Let* $E\subseteq\Delta_0\in\mathscr{D}$ *be a Borel set*, $M>1$, *and* $\delta>0$, *and set*
$$\mathscr{P}=\{\Delta\subseteq\Delta_0:\Delta\cap E\neq\varnothing,\ \textit{there exists } \xi\in B_X(\zeta_\Delta,M\ell(\Delta))\textit{ such that }\operatorname{dist}(\xi,E)\geq\delta\ell(\Delta)\}.$$
Then there is $C_1=C_1(M,\delta,C_\mu)>0$ *so that, for all* $\Delta'\subseteq\Delta_0$,
$$\sum_{\substack{\Delta\subseteq\Delta'\\ \Delta\in\mathscr{P}}}\mu(\Delta)\leq C_1\mu(\Delta'). \tag{3-2}$$

The theorem is stated in [Azzam 2014] in slightly more generality. For the reader's convenience, we provide a shorter proof in the Appendix.

Let $M,\delta>0$, to be decided later, and let $\mathscr{P}$ be the set from Lemma 3.4 applied to our set E. Our goal now is to construct a metric space Y containing X, then a curve $\Gamma\subseteq Y$ that contains E as a subset, and then show it has finite length. We will do this by adding bridges through Y between net points around cubes in $\mathscr{P}$ since these are the cubes where E has large holes and thus potentially has big gaps or disconnections. We don't need the endpoints of these bridges to be in E, but their union plus the set E will be connected. We now proceed with the details.

Let $\widetilde{X}=\bigcup X_n$, and equip $C(X)\oplus\mathbb{R}^{\widetilde{X}\times\widetilde{X}}$ (where $\mathbb{R}^{\widetilde{X}\times\widetilde{X}}=\prod_{\alpha\in\widetilde{X}\times\widetilde{X}}\mathbb{R}$; see [Munkres 1975, p. 112–117] for the notation) with norm $|a\oplus b|=\max\{|a|,|b|\}$, where the norm on $\mathbb{R}^{\widetilde{X}\times\widetilde{X}}$ is the ℓ^2 norm.

For $x,y\in\widetilde{X}$, let $[x,y]$ denote the straight line segment between them in $C(X)\oplus\mathbb{R}^{\widetilde{X}\times\widetilde{X}}$, $e_{(x,y)}$ is the unit vector corresponding to the (x,y) coordinate in $\mathbb{R}^{\widetilde{X}\times\widetilde{X}}$, and define
$$\begin{aligned}[x,y]^*:=&\big[x,(x,|x-y|e_{(x,y)})\big]\cup\big[y,(y,|x-y|e_{(x,y)})\big]\cup\big[(x,|x-y|e_{(x,y)}),(y,|x-y|e_{(x,y)})\big]\\ &\subseteq C(X)\oplus\mathbb{R}^{\widetilde{X}\times\widetilde{X}}.\end{aligned}$$

The set $[x,y]^*$ is two segments going straight up from x and y, respectively, in the $e_{(x,y)}$ direction and a segment connecting the endpoints, thus giving a polygonal curve connecting x to y that hops out of $C(X)$. Let

$$Y = X \cup \bigcup_{x,y \in \widetilde{X}} [x, y]^*,$$

and define a metric on Y (also denoted by $|\cdot|$) by setting

$$|x - y| = \inf \sum_{i=1}^{N} |x_i - x_{i+1}|$$

where $x_1 = x$, $x_{N+1} = y$, and for each i, $\{x_i, x_{i+1}\} \subseteq X$ or $\{x_i, x_{i+1}\} \subseteq [x', y']^*$ for some $x', y' \in \widetilde{X}$. It is easy to check that the resulting metric space Y is separable and X is a metric subspace in Y. Moreover, the following lemma is immediate from the definition of Y.

Lemma 3.5. *Let $F \subseteq X$ be compact and $x, y \in \widetilde{X}$. Then*

$$\operatorname{dist}([x, y]^*, F) = \operatorname{dist}(\{x, y\}, F).$$

We will let

$$B_\Delta := B_Y(\zeta_\Delta, \ell(\Delta)) \supseteq B_X(\zeta_\Delta, \ell(\Delta)).$$

For $\Delta \in \mathscr{D}_n$, let

$$\Gamma_\Delta = \bigcup \left\{ [x, y]^* \subseteq C(X) \oplus \mathbb{R}^{\widetilde{X} \times \widetilde{X}} : x, y \in X_{n+n_0} \cap M B_\Delta \right\}$$

where n_0 is an integer we will pick later. Note that Γ_Δ is connected and contains ζ_Δ.

Now define

$$\Gamma = E \cup \bigcup_{\Delta \in \mathscr{P}} \Gamma_\Delta.$$

Lemma 3.6. $$\mathscr{H}^1(\Gamma) < \infty.$$

Proof. We first claim that

$$\mathscr{H}^1(E) \leq 10\mu(E). \tag{3-3}$$

Indeed, let $0 < \delta < r_0$. Take any countable collection of balls centered on E of radii less than δ that cover E. Since μ is doubling, we can use the Vitali covering theorem [Heinonen 2001, Theorem 1.2] to find a countable subcollection of disjoint balls B_i with radii $r_i < \delta$ centered on E so that $E \subseteq \bigcup 5B_i$. Then

$$\mathscr{H}^1_\delta(E) \leq \sum 10r_i \leq 10 \sum \mu(B_i) \leq 10\mu(\{x \in X : \operatorname{dist}(x, E) < \delta\}).$$

Since $\bigcap_{\delta > 0} \{x \in X : \operatorname{dist}(x, E) < \delta\} = E$, sending $\delta \to 0$, we obtain $\mathscr{H}^1(E) \leq 10\mu(E)$, which proves the claim.

With this estimate in hand, we have

$$\begin{aligned}\mathscr{H}^1(\Gamma) &\leq \mathscr{H}^1(E)+\sum_{\Delta\in\mathscr{P}}\mathscr{H}^1(\Gamma_\Delta) \overset{(3\text{-}3)}{\leq} 10\mu(E)+C\sum_{\Delta\in\mathscr{P}}\ell(\Delta)\\ &\overset{(3\text{-}1)}{\leq} 10\mu(E)+C\sum_{\Delta\in\mathscr{P}}\mu(\Delta) \overset{(3\text{-}2)}{\leq} 10\mu(E)+C\mu(\Delta_0)<\infty\end{aligned}$$

where C here stands for various constants that depend only on δ, M, n_0, ρ, and the doubling constant C_μ. □

Lemma 3.7. Γ *is compact.*

Proof. To see this, let $x_n \in \Gamma$ be any sequence. If $x_n \in \Gamma_\Delta$ infinitely many times for some $\Delta \in \mathscr{P}$ or is in E infinitely many times, then since each of these sets are compact, we can find a convergent subsequence with a limit in Γ. Otherwise, x_n visits infinitely many Γ_Δ. Let x_{n_j} be a subsequence so that $x_{n_j} \in \Gamma_{\Delta_j}$ where each $\Delta_j \in \mathscr{P}$ is distinct. Then $\ell(\Delta_j) \to 0$, and since $\Delta \cap E \neq \varnothing$ for all $\Delta \in \mathscr{P}$, $\operatorname{dist}(x_{n_j}, E) \to 0$. Pick $x'_{n_j} \in E \cap \Delta_j$. Since E is compact, there is a subsequence $x'_{n_{j_k}}$ converging to a point in E, and $x_{n_{j_k}}$ will have the same limit. We have thus shown that any sequence in Γ has a convergent subsequence, which implies Γ is compact. □

Lemma 3.8. *A compact connected metric space X of finite length can be parametrized by a Lipschitz image of an interval in $\mathbb{R}$; that is, $X = f([0, 1])$ where $f : [0, 1] \to X$ is Lipschitz.*

A proof of this fact for Hilbert spaces is given in [Schul 2007, Corollary 3.7], but the same proof works in our setting, so we omit it. Hence, to show that Γ (and hence E) is rectifiable, all that remains to show is that Γ is connected.

Lemma 3.9. *The set Γ is connected.*

Proof. Suppose for the sake of a contradiction that there exist two open and disjoint sets A and B that cover Γ, and set $\Gamma_A = \Gamma \cap A$ and $\Gamma_B = \Gamma \cap B$. Suppose without loss of generality that $\Gamma_{\Delta_0} \subseteq \Gamma_A$, which we may do since Γ_{Δ_0} is connected. We sort the proof into a series of steps.

(a) $\Gamma_B \subseteq 2B_{\Delta_0}$. To see this, suppose instead that there is $z \in \Gamma_B \setminus 2B_{\Delta_0}$. Then $z \in [x, y]^* \subseteq \Gamma_\Delta$ for some $\Delta \in \mathscr{P}$. Moreover, $\operatorname{dist}(z, \{x, y\}) \leq 2|x-y| \leq 4M\ell(\Delta)$ since $x, y \in MB_\Delta$. Since $\zeta_\Delta \in \Delta \subseteq \Delta_0$ and $x \in MB_\Delta$, we get

$$\begin{aligned}\ell(\Delta_0) &\leq \operatorname{dist}(z, B_{\Delta_0}) \leq |z-x| + \operatorname{dist}(x, B_{\Delta_0}) \leq 4M\ell(\Delta)+M\ell(\Delta)\\ &= 5M\ell(\Delta).\end{aligned}$$

For n_0 large enough so that $5M\rho^{n_0} < 1$, this implies $\zeta_\Delta \in X_{n+n_0} \cap MB_{\Delta_0}$ and so $\Gamma_\Delta \cap \Gamma_{\Delta_0} \neq \varnothing$. Hence, $\Gamma_\Delta \subseteq \Gamma_A$ since Γ_Δ is connected, contradicting that $z \in \Gamma_B$. This proves the claim.

(b) The open sets $A' = A \cup (\overline{4B_{\Delta_0}})^c$ and $B' = B \cap 2B_{\Delta_0}$ are disjoint and cover Γ. First, observe that

$$\begin{aligned}A' \cap B' &= (A \cap B \cap 2B_{\Delta_0}) \cup ((\overline{4B_{\Delta_0}})^c \cap B \cap 2B_{\Delta_0})\\ &\subseteq (A \cap B) \cup ((\overline{4B_{\Delta_0}})^c \cap 2B_{\Delta_0}) = \varnothing.\end{aligned}$$

Moreover, by part (a),

$$\Gamma \cap (A' \cup B') \supseteq \Gamma_A \cup (\Gamma_B \cap 2B_{\Delta_0}) = \Gamma_A \cup \Gamma_B = \Gamma,$$

which completes the proof of this step.

(c) Set $\Gamma_{A'} = \Gamma \cap A'$ and $\Gamma_{B'} = \Gamma \cap B'$. These sets are disjoint by part (b), and hence, they are compact since Γ was compact. We define new open sets

$$A'' = (\overline{4B_{\Delta_0}})^c \cup \bigcup_{\xi \in \Gamma_{A'}} B_Y(\xi, \operatorname{dist}(\xi, \Gamma_{B'})/2)$$

and

$$B'' = \bigcup_{\xi \in \Gamma_{B'}} B_Y(\xi, \operatorname{dist}(\xi, \Gamma_{A'})/2).$$

We claim these sets are disjoint. Suppose there is $z \in A'' \cap B''$. Then $z \in B_Y(\xi, \operatorname{dist}(\xi, \Gamma_{A'})/2)$ for some $\xi \in \Gamma_{B'}$. If we also have $z \in B_Y(\xi', \operatorname{dist}(\xi', \Gamma_{B'})/2)$ for some $\xi' \in \Gamma_{A'}$, then

$$\max\{\operatorname{dist}(\xi, \Gamma_{B'}), \operatorname{dist}(\xi', \Gamma_{A'})\} \le |\xi - \xi'| \le |\xi - z| + |z - \xi| < \frac{\operatorname{dist}(\xi, \Gamma_{B'})}{2} + \frac{\operatorname{dist}(\xi', \Gamma_{A'})}{2},$$

which is a contradiction, so we must have $z \in (\overline{4B_{\Delta_0}})^c$. Since $\xi \in \Gamma_{B'}$, we know $\xi \in 2B_{\Delta_0}$ by part (a), and $\zeta_{\Delta_0} \in \Gamma_{\Delta_0} \subseteq \Gamma_{A'}$ implies $\operatorname{dist}(\xi, \Gamma_{A'}) \le 2\ell(\Delta_0)$. Hence,

$$B_Y(\xi, \operatorname{dist}(\xi, \Gamma_{A'})/2) \subseteq B_Y(\xi, \ell(\Delta_0)) \subseteq B_Y(\zeta_{\Delta_0}, 3\ell(\Delta_0)) = 3B_{\Delta_0},$$

which proves the claim.

(d) Note that $X \setminus (A'' \cup B'')$ is nonempty since X is connected and A'' and B'' are disjoint open sets. Moreover, $X \setminus (A'' \cup B'') \subseteq \overline{4B_{\Delta_0}}$ and hence a bounded set; since X is a doubling metric space, $X \setminus (A'' \cup B'')$ is in fact totally bounded and thus compact by the Heine–Borel theorem. This implies we can find a point

$$z \in X \setminus (A'' \cup B'') \subseteq \overline{4B_{\Delta_0}}$$

of maximal distance from the compact set Γ.

(e) Let $\xi \in E$ be the closest point to z and Δ the smallest cube containing ξ so that $z \in 5B_\Delta$; since $z \in \overline{4B_{\Delta_0}} \subseteq 5B_{\Delta_0}$, this is well defined. We claim $\Delta \in \mathscr{P}$. If Δ_1 denotes the child of Δ that contains ξ, then $z \notin 5B_{\Delta_1}$, and so

$$\begin{aligned}\operatorname{dist}(z, E) = |\xi - z| &\ge |z - \zeta_{\Delta_1}| - |\zeta_{\Delta_1} - \xi| \ge 5\ell(\Delta_1) - \ell(\Delta_1) \\ &= 4\rho\ell(\Delta).\end{aligned} \tag{3-4}$$

Thus, for $M > 10$, $B_X(z, 4\rho\ell(\Delta)) \subseteq MB_\Delta \setminus E$, so if $\delta < 4\rho$, then $\Delta \in \mathscr{P}$, which proves the claim.

(f) Since $\Delta \in \mathscr{P}$, $X_{n(\Delta)+n_0}$ is a maximal $\rho^{n(\Delta)+n_0}$-net,

$$\rho^{n(\Delta)+n_0} < \rho^{n_0}\ell(\Delta) < \ell(\Delta),$$

and $z \in 5B_\Delta$, we can find

$$\begin{aligned} \zeta &\in X_{n(\Delta)+n_0} \cap B_X(z, \rho^{n(\Delta)+n_0}) && (3\text{-}5)\\ &\subseteq X_{n(\Delta)+n_0} \cap B_X(\zeta_\Delta, 5\ell(\Delta)+\rho^{n(\Delta)+n_0}) \\ &\subseteq X_{n(\Delta)+n_0} \cap B_X(\zeta_\Delta, 6\ell(\Delta)) \subseteq \Gamma_\Delta, && (3\text{-}6) \end{aligned}$$

where the last containment follows if we assume $M > 6$.

Since Γ_Δ is connected and A' and B' are disjoint open sets, we may without loss of generality suppose $\Gamma_{A'} \supseteq \Gamma_\Delta$ and let $\zeta' \in \Gamma_{B'}$ be the closest point to ζ. Then

$$|z-\zeta| \geq |\zeta-\zeta'|/2 = \operatorname{dist}(\zeta, \Gamma_{B'})/2 \tag{3-7}$$

since otherwise would imply $z \in B_Y(\zeta, \operatorname{dist}(\zeta, \Gamma_{B'})/2) \subseteq A''$, contradicting that $z \in X \setminus (A'' \cup B'')$.

We may assume $\zeta' \in \Gamma_{\Delta'}$ for some $\Delta' \in \mathscr{P}$, and we assume Δ' is the largest such cube for which this happens. Note that this implies $\Gamma_{\Delta'} \subseteq \Gamma_{B'}$ since $\zeta' \in \Gamma_{B'} \cap \Gamma_{\Delta'}$ and $\Gamma_{\Delta'}$ is connected. By Lemma 3.5 with $F = \{\zeta\}$, we can assume $\zeta' \in X$, and so $\zeta' \in X_{n(\Delta')+n_0} \cap MB_{\Delta'}$.

(g) We claim that $n(\Delta)+1 \leq n(\Delta') \leq n(\Delta)+2$. Note that, since

$$5\rho^{n(\Delta)+n_0} \leq \ell(\Delta)\rho^{n_0} \leq \rho\ell(\Delta) < \ell(\Delta), \tag{3-8}$$

we have

$$|\zeta'-\zeta_\Delta| \;\leq\; |\zeta'-\zeta|+|\zeta-\zeta_\Delta| \overset{(3\text{-}6)}{\underset{}{\overset{(3\text{-}7)}{<}}} 2|\zeta-z|+6\ell(\Delta) \overset{(3\text{-}5)}{<} 2\rho^{n(\Delta)+n_0}+6\ell(\Delta) \overset{(3\text{-}8)}{\leq} 8\ell(\Delta). \tag{3-9}$$

Thus, for $M > 8$, we must have $n(\Delta') > n(\Delta)$; otherwise, since $\xi \in \Delta \subseteq B_\Delta$, we would have

$$\zeta' \in X_{n(\Delta')+n_0} \cap 8B_\Delta \subseteq X_{n(\Delta)+n_0} \cap MB_\Delta \subseteq \Gamma_\Delta$$

so that $\Gamma_\Delta \cap \Gamma_{\Delta'} \neq \varnothing$, which implies $\Gamma_{A'} \cap \Gamma_{B'} \neq \varnothing$, a contradiction. Thus, $\ell(\Delta') < \ell(\Delta)$, which proves the first inequality in the claim.

Note this implies $\ell(\Delta') \leq \rho\ell(\Delta)$. Let $\xi' \in \Delta' \cap E$ (which exists since $\Delta' \in \mathscr{P}$). Since $\zeta' \in MB_{\Delta'}$,

$$\begin{aligned} 4\rho\ell(\Delta) &\overset{(3\text{-}4)}{\leq} \operatorname{dist}(z, E) \leq |\xi'-z| \leq |\xi'-\zeta_{\Delta'}|+|\zeta_{\Delta'}-\zeta'|+|\zeta'-\zeta|+|\zeta-z| \\ &\overset{(3\text{-}7)}{\leq} \ell(\Delta')+M\ell(\Delta')+2|\zeta-z|+|\zeta-z| \leq (M+1)\ell(\Delta')+3\rho^{n(\Delta)+n_0} \\ &\overset{(3\text{-}8)}{\leq} (M+1)\ell(\Delta')+\rho\ell(\Delta) \end{aligned}$$

and so

$$\frac{3\rho}{M+1}\ell(\Delta) \leq \ell(\Delta').$$

Thus, $\rho < 3/(M+1)$ implies $\rho^2\ell(\Delta) \leq \ell(\Delta')$, and so $n(\Delta') \leq n(\Delta)+2$, which finishes the claim.

(h) Now we'll show that $\Gamma_\Delta \cap \Gamma_{\Delta'} \neq \varnothing$. Observe that

$$|\zeta_\Delta-\zeta_{\Delta'}| \leq |\zeta_\Delta-\zeta'|+|\zeta'-\zeta_{\Delta'}| \overset{(3\text{-}9)}{\leq} 8\ell(\Delta)+M\ell(\Delta') \leq (8+M\rho)\ell(\Delta) < M\ell(\Delta) \tag{3-10}$$

if $\rho^{-1} > M > 9$. Since $n(\Delta') \le n(\Delta)+2$, we have that $\zeta_{\Delta'} \in X_{n(\Delta)+n_0} \cap MB_\Delta$ for $n_0 \ge 2$ and so $\zeta_{\Delta'} \in \Gamma_\Delta$. But $\zeta_{\Delta'} \in X_{n(\Delta')+n_0} \cap MB_{\Delta'} \subseteq \Gamma_{\Delta'}$; thus, $\Gamma_\Delta \cap \Gamma_{\Delta'} \neq \varnothing$, which proves the claim.

This gives us a grand contradiction since $\Gamma_\Delta \subseteq \Gamma_{A'}$ and $\Gamma_{\Delta'} \subseteq \Gamma_{B'}$, and we assumed these sets to be disjoint. □

Combining Lemmas 3.6, 3.7, 3.8, and 3.9, we have now shown that E is contained in the Lipschitz image of an interval in $\mathbb{R}$. This completes the proof of Lemma 3.2. □

Appendix: Proof of Lemma 3.4

For $\Delta \in \mathscr{D}$, define $B_\Delta = B_X(\zeta_\Delta, \ell(\Delta))$. For $\Delta \in \mathscr{P}$, let $\xi_\Delta \in MB_\Delta$ be such that $\operatorname{dist}(\xi, E) \ge \delta\ell(\Delta)$. Let $\mathscr{M}$ be the collection of maximal cubes for which $2B_\Delta \subseteq E^c$ and $\tilde{\Delta} \in \mathscr{M}$ be the largest cube containing ξ_Δ. Then if $\tilde{\Delta}^1$ denotes the parent cube of $\tilde{\Delta}$, $2B_{\tilde{\Delta}^1} \cap E \neq \varnothing$, and so

$$\delta\ell(\Delta) \le \operatorname{dist}(\xi_\Delta, E) \le \operatorname{diam} 2B_{\tilde{\Delta}^1} \le 4\ell(\tilde{\Delta}^1) = \frac{4}{\rho}\ell(\tilde{\Delta}). \tag{A-1}$$

Moreover,

$$\ell(\tilde{\Delta}) \le \frac{2M}{c_0}\ell(\Delta), \tag{A-2}$$

for otherwise $\tilde{\Delta} \supseteq c_0 B_{\tilde{\Delta}} \supseteq MB_\Delta \supseteq \Delta$, and since $\Delta \cap E \neq \varnothing$, this means $2B_{\tilde{\Delta}} \cap E \neq \varnothing$, contradicting our definition of $\tilde{\Delta}$.

Let N_Δ be such that

$$2^{N_\Delta} c_0 \ell(\tilde{\Delta}) > 2M\ell(\Delta) > 2^{N_\Delta - 1} c_0 \ell(\tilde{\Delta}). \tag{A-3}$$

Then $2^{N_\Delta} c_0 B_{\tilde{\Delta}} \supseteq MB_\Delta$, and $2^{N_\Delta} < \dfrac{4M\ell(\Delta)}{c_0\ell(\tilde{\Delta})}$, so

$$N_\Delta < \log_2\left(\frac{4M\ell(\Delta)}{c_0\ell(\tilde{\Delta})}\right). \tag{A-4}$$

Thus,

$$\begin{aligned} \frac{\mu(\tilde{\Delta})}{\mu(\Delta)} &\ge \frac{\mu(c_0 B_{\tilde{\Delta}})}{\mu(\Delta)} \overset{\text{(1-1)}}{\ge} \frac{\mu(2^{N_\Delta} c_0 B_{\tilde{\Delta}})}{C_\mu^{N_\Delta}\mu(\Delta)} \overset{\text{(A-3)}}{\ge} \frac{\mu(MB_\Delta)}{C_\mu^{N_\Delta}\mu(\Delta)} \\ &\overset{\text{(A-4)}}{\ge} C_\mu^{\log_2 c_0/(4M)} \left(\frac{\ell(\tilde{\Delta})}{\ell(\Delta)}\right)^{\log_2 C_\mu} \overset{\text{(A-1)}}{\ge} C_\mu^{\log_2 c_0/(4M)} \left(\frac{4}{\rho}\right)^{\log_2 C_\mu} =: a. \end{aligned} \tag{A-5}$$

Since μ is doubling and Δ and Δ' are always of comparable sizes by (A-1) and (A-2), there is b depending on M, δ, ρ, c_0, and C_μ such that at most b many cubes $\Delta \in \mathscr{M}$ with $\tilde{\Delta} = \Delta'$ for some fixed Δ'. Hence, for $\Delta' \subseteq \Delta_0$ with $\Delta \cap E \neq \varnothing$,

$$\begin{aligned} \sum_{\substack{\Delta \subseteq \Delta' \\ \Delta \in \mathscr{P}}} \mu(\Delta) &\overset{\text{(A-5)}}{\le} \sum_{\substack{\Delta \subseteq \Delta' \\ \Delta \in \mathscr{P}}} a\mu(\tilde{\Delta}) = \sum_{\substack{\Delta' \in \mathscr{M} \\ \Delta \subseteq MB_{\Delta_0}}} \sum_{\substack{\Delta \subseteq \Delta' \\ \Delta \in \mathscr{P} \\ \tilde{\Delta} = \Delta'}} a\mu(\tilde{\Delta}) \le \sum_{\substack{\Delta' \in \mathscr{M} \\ \Delta \subseteq MB_{\Delta_0}}} ab\mu(\Delta') \\ &\le ab\mu(MB_{\Delta_0} \setminus E) \le ab\mu(MB_{\Delta_0}) \overset{\text{(1-1)}}{\le} abC_\mu^{\log_2 M/c_0 + 1}\mu(c_0 B_{\Delta_0}) \le abC_\mu^{\log_2 M/c_0 + 1}\mu(\Delta_0). \end{aligned}$$

This finishes the proof of Lemma 3.4.

Acknowledgements

The authors thank Raanan Schul for his encouragement and helpful discussions, which improved the result, as well as John Garnett and the anonymous referee whose advice greatly improved the readability of the paper.

References

[Azzam 2014] J. Azzam, "Sets of absolute continuity for harmonic measure in NTA domains", preprint, 2014. arXiv 1410.2782v2

[Azzam et al. 2015] J. Azzam, G. David, and T. Toro, "Wasserstein distance and the rectifiability of doubling measures, I", *Math. Ann.* (online publication April 2015).

[Badger and Schul 2015a] M. Badger and R. Schul, "Multiscale analysis of 1-rectifiable measures: necessary conditions", *Math. Ann.* **361**:3–4 (2015), 1055–1072. MR 3319560 Zbl 1314.28003

[Badger and Schul 2015b] M. Badger and R. Schul, "Two sufficient conditions for rectifiable measures", *Proc. Amer. Math. Soc.* (online publication October 2015).

[Christ 1990] M. Christ, "A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral", *Colloq. Math.* **60–61**:2 (1990), 601–628. MR 92k:42020 Zbl 0758.42009

[David 1988] G. David, "Morceaux de graphes lipschitziens et intégrales singulières sur une surface", *Rev. Mat. Iberoamericana* **4**:1 (1988), 73–114. MR 90h:42026 Zbl 0696.42011

[Federer 1969] H. Federer, *Geometric measure theory*, Grundl. Math. Wissen. **153**, Springer, New York, 1969. MR 41 #1976 Zbl 0176.00801

[Garnett et al. 2010] J. Garnett, R. Killip, and R. Schul, "A doubling measure on $\mathbb{R}^d$ can charge a rectifiable curve", *Proc. Amer. Math. Soc.* **138**:5 (2010), 1673–1679. MR 2011a:28018 Zbl 1196.28007

[Heinonen 2001] J. Heinonen, *Lectures on analysis on metric spaces*, Springer, New York, 2001. MR 2002c:30028 Zbl 0985.46008

[Hytönen and Martikainen 2012] T. Hytönen and H. Martikainen, "Non-homogeneous Tb theorem and random dyadic cubes on metric measure spaces", *J. Geom. Anal.* **22**:4 (2012), 1071–1107. MR 2965363 Zbl 1261.42017

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, Cambridge Studies Advanced Math. **44**, Cambridge University, 1995. MR 96h:28006 Zbl 0819.28004

[Munkres 1975] J. R. Munkres, *Topology: a first course*, Prentice Hall, Englewood Cliffs, NJ, 1975. MR 57 #4063 Zbl 0306.54001

[Schul 2007] R. Schul, "Subsets of rectifiable curves in Hilbert space — the analyst's TSP", *J. Anal. Math.* **103** (2007), 331–375. MR 2008m:49205 Zbl 1152.28006

Received 14 Jan 2015. Revised 22 Jun 2015. Accepted 11 Oct 2015.

JONAS AZZAM: jazzam@mat.uab.cat
Departament de Matemàtiques, Universitat Autònoma de Barcelona, Edifici C Facultat de Ciències, 08193 Bellaterra, Spain

MIHALIS MOURGOGLOU: mmourgoglou@crm.cat
Departament de Matemàtiques, Universitat Autònoma de Barcelona and Centre de Recerca Matemàtica, Edifici C Facultat de Ciències, 08193 Bellaterra, Spain

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.111

CONSTRUCTION OF HADAMARD STATES BY CHARACTERISTIC CAUCHY PROBLEM

CHRISTIAN GÉRARD AND MICHAŁ WROCHNA

We construct Hadamard states for Klein–Gordon fields in a spacetime M_0 equal to the interior of the future lightcone C from a base point p in a globally hyperbolic spacetime (M, g).

Under some regularity conditions at the future infinity of C, we identify a boundary symplectic space of functions on C, which allows us to construct states for Klein–Gordon quantum fields in M_0 from states on the CCR algebra associated to the boundary symplectic space. We formulate the natural microlocal condition on the boundary state on C, ensuring that the bulk state it induces in M_0 satisfies the Hadamard condition.

Using pseudodifferential calculus on the cone C, we construct a large class of Hadamard states on the boundary with pseudodifferential covariances and characterize the pure states among them. We then show that these pure boundary states induce pure Hadamard states in M_0.

1. Introduction

Hadamard states are widely accepted as physically admissible states for noninteracting quantum fields on a curved spacetime, one of the main reasons being their link with the renormalization of the stress–energy tensor, a basic step in the formulation of semiclassical Einstein equations. Furthermore, they are nowadays considered a necessary ingredient in the perturbative formulation of interacting (nonlinear) theories (see the recent review articles [Khavkine and Moretti 2015; Hollands and Wald 2015]).

For Klein–Gordon fields, the construction of Hadamard states amounts to finding bisolutions of the Klein–Gordon equation (called in this context *two-point functions* and denoted here by $\lambda^{\pm}$) with a specified wavefront set (that is, verifying the *microlocal spectrum condition*) and satisfying additionally a positivity property [Radzikowski 1996].

There exist several ways to construct Hadamard states for Klein–Gordon fields: the first method relies on the Fulling–Narcowich–Wald deformation argument [Fulling et al. 1981], which reduces the construction of Hadamard states on an arbitrary spacetime to the case of ultrastatic spacetimes, where vacuum or thermal states are easily shown to be Hadamard states.

The second approach, worked out in [Junker 1995; Junker and Schrohe 2002; Gérard and Wrochna 2014], uses pseudodifferential calculus on a fixed Cauchy surface Σ in (M, g) and relies on the construction of a parametrix for the Cauchy problem on Σ. To use pseudodifferential calculus, some restrictions on Σ and on the behavior of the metric g at spatial infinity are necessary. On the other hand, the method

MSC2010: 35S05, 81T20.
Keywords: Hadamard states, microlocal spectrum condition, pseudodifferential calculus, characteristic Cauchy problem, curved spacetimes.

produces a large classes of rather explicit Hadamard states, whose covariances, expressed in terms of Cauchy data, are pseudodifferential operators.

Another method, initiated by Moretti [2006; 2008] applies to *conformal field equations*, like the conformal wave equation, on an *asymptotically flat vacuum spacetime* (M_0, g_0). By asymptotic flatness, there exists a metric $\tilde{g}_0$, conformal to g_0, and a spacetime $(M, \tilde{g})$ such that $(M_0, \tilde{g}_0)$ can be causally embedded as an open set in $(M, \tilde{g})$ with the boundary $C = \partial M_0$ of M_0 being *null* in $(M, \tilde{g})$. States on the *boundary symplectic space*, containing the traces on C of solutions of the wave equation in M_0, naturally induce states inside M_0.

This method has been successfully applied in [Moretti 2006; 2008] to construct a distinguished Hadamard state for asymptotically flat vacuum spacetimes with past time infinity and then extended to several other geometrical situations in [Dappiaggi et al. 2009; 2011; Brum and Jorás 2015]. Further results also include generalizations to Maxwell fields [Dappiaggi and Siemssen 2013] and linearized gravity [Benini et al. 2014].

In the present paper we rework the above strategy systematically in terms of the associated *characteristic Cauchy problem* in order to construct a large class of Hadamard states (instead of a preferred single one) and to characterize the pure ones. For the sake of clarity, we do not impose geometrical assumptions on M_0 that allow one to correctly embed it in a larger spacetime M.

Instead we go the other way around and work in an a priori arbitrary globally hyperbolic spacetime (M, g), fix a base point p and consider the interior of the future lightcone

$$C := \partial J^+(p)\backslash\{p\}$$

as the spacetime M_0 of main interest, that is, $M_0 := I^+(p)$, where $I^+(p)$ (resp. $J^+(p)$) is the timelike (resp. causal) shadow of p; see [Wald 1984, Section 8.1].

We make the following assumption on the geometry of C.

Hypothesis 1.1. We assume that there exists $f \in C^\infty(M)$ such that:

(1) $C \subset f^{-1}(\{0\})$, $\nabla_a f \neq 0$ on C, $\nabla_a f(p) = 0$ and $\nabla_a \nabla_b f(p) = -2 g_{ab}(p)$.

(2) The vector field $\nabla^a f$ is complete on C.

Using Hypothesis 1.1 one can construct coordinates (f, s, θ) near C such that $C \subset \{f = 0\}$ and

$$g\restriction_C = -2\, df\, ds + h(s, \theta)\, d\theta^2,$$

where $h(s, \theta)\, d\theta^2$ is a Riemannian metric on $\mathbb{S}^{d-1}$.

This choice of coordinates allows one to identify C with $\widetilde{C} := \mathbb{R} \times \mathbb{S}^{d-1}$. A natural space of smooth functions on $\widetilde{C}$ is then provided by $\mathcal{H}(\widetilde{C})$ — the intersection of Sobolev spaces of all orders, defined using the standard metric $m(\theta)\, d\theta^2$ on $\mathbb{S}^{d-1}$.

We consider the Klein–Gordon operator $P = -\Box_g + r(x)$ (with $r(x) \in C^\infty(M)$ real-valued) and its restriction on M_0, denoted by $P_0 := P\restriction_{M_0}$. The *bulk-to-boundary correspondence* can be expressed in this setup as follows. For an appropriate choice of $\beta(s, \theta) \in C^\infty(M_0)$, the restriction map

$$\rho\phi := (\beta^{-1}\phi)\restriction_C, \quad \phi \in C^\infty_{\mathrm{sc}}(M_0),$$

is a monomorphism[1] between the symplectic space of smooth, space-compact solutions of P_0 (endowed with the usual symplectic form induced by the causal propagator) and $\mathscr{H}(\widetilde{C})$, equipped with the symplectic form

$$\bar{g}_1 \sigma_C g_2 := \int_{\mathbb{R}\times\mathbb{S}^{d-1}} (\partial_s \bar{g}_1 g_2 - \bar{g}_1\, \partial_s g_2)|m|^{1/2}(\theta)\, ds\, d\theta, \quad g_1, g_2 \in \mathscr{H}(\widetilde{C}). \tag{1-1}$$

Thus, a quasifree state on $(\mathscr{H}(\widetilde{C}), \sigma_C)$ with two-point functions $\lambda^{\pm}$ induces a unique quasifree state on the usual symplectic space associated to P_0.

Product-type pseudodifferential operators. In [Gérard and Wrochna 2014] we constructed Hadamard states whose two-point functions on a Cauchy surface Σ are pseudodifferential operators. In the present case, the obvious difference is that on the cone C the coordinate s is distinguished both from the point of view of the microlocal spectrum condition (from now on abbreviated (μsc)) and in the expression (1-1) for the symplectic form. This suggests that one should rather consider *product-type* pseudodifferential operators $\Psi^{p_1,p_2}(\widetilde{C})$ with symbols satisfying estimates

$$|\partial_s^{\alpha_1}\, \partial_\sigma^{\beta_1}\, \partial_\theta^{\alpha_2}\, \partial_\eta^{\beta_2} a(s, \theta, \sigma, \eta)| \in O(\langle\sigma\rangle^{p_1-|\beta_1|}\langle\eta\rangle^{p_2-|\beta_2|})$$

in the covariables $\xi = (\sigma, \eta)$ relative to the decomposition $\widetilde{C} = \mathbb{R}\times\mathbb{S}^{d-1}$. Actually, to cope with the issue that σ_C is defined using an operator $D_s := \mathrm{i}^{-1}\partial_s$ whose spectrum is not separated from $\{0\}$ (analogously to the infrared problem in massless theories), we need to introduce a larger class $\widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$ that includes some operators whose symbol is discontinuous at $\eta = 0$. Namely, we set

$$\widetilde{\Psi}^{p_1,p_2}(\widetilde{C}) := \Psi^{p_1,p_2}(\widetilde{C}) + B^{-\infty}\Psi^{p_2}(\widetilde{C}),$$

where $B^{-\infty}\Psi^{p_2}(\widetilde{C})$ is the class of pseudodifferential operators of order p_2 (in the θ variables) with values in operators on $\mathbb{R}$ that infinitely increase Sobolev regularity. Then, for instance, $|D_s| \otimes \mathbb{1}_\theta \in \widetilde{\Psi}^{1,0}(\widetilde{C})$ although it is not in the pseudodifferential class $\Psi^{1,0}(\widetilde{C})$.

Summary of results. Our main results can be summarized as follows. We always assume Hypothesis 1.1. If E and F are topological vector spaces, we write $T : E \to F$ to mean $T : E \to F$ is linear and continuous.

(1) For pairs[2] of two-point functions $\lambda^{\pm}$ on C satisfying $\lambda^{\pm} : \mathscr{H}(C) \to \mathscr{H}(C)$, we give in Theorem 5.3 conditions on $\mathrm{WF}(\lambda^{\pm})$ that guarantee that the corresponding two-point functions on M_0 satisfy (μsc). This is essentially an adaptation of the results of [Moretti 2008] to our framework.

(2) In Theorem 7.4 we construct a large class of Hadamard states by specifying their two-point functions $\lambda^{\pm} \in \widetilde{\Psi}^{0,0}(\widetilde{C})$ on the cone.

(3) In Theorem 8.2 we characterize the subclass of Hadamard states constructed in (2), which additionally are pure on the symplectic space $(\mathscr{H}(\widetilde{C}), \sigma_C)$ on the cone. It turns out that they can be parametrized by a single operator in $\widetilde{\Psi}^{-\infty,0}(\widetilde{C})$.

[1]By monomorphism of symplectic spaces we mean an injective linear map that intertwines the symplectic forms.

[2]We work with charged fields, in which case it is natural to associate a *pair* of two-point functions to a quasifree state; see Section 3B1. The charged and neutral approaches are equivalent.

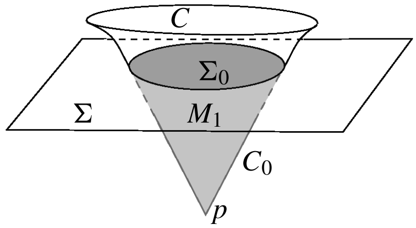

Figure 1. The Cauchy surface Σ in the future of p.

(4) In Theorem 8.4 we prove that if $\dim M \geq 4$ then the pure states considered in (3) induce pure states in the interior M_0 of the cone.

In Section 2C we argue that Hypothesis 1.1 covers the case when M_0 is an asymptotically flat vacuum spacetime with future time infinity, after a conformal transformation. Thus, our result (4) solves an open question of Moretti [2008] for $\dim M \geq 4$.

Characteristic Cauchy problem. The proof of our main result (4) relies on rather standard results on the *characteristic Cauchy problem* (also called the *Goursat problem* in the literature) in appropriate Sobolev spaces.

Let Σ be a Cauchy surface for (M, g) in the future of $\{p\}$ and $\Sigma_0 := \Sigma \cap M_0$. We set

$$M_1 := I^-(\Sigma_0; M) \cap M_0 \quad \text{and} \quad C_0 := (J^-(\Sigma_0; M) \cap C) \cup \{p\};$$

see Figure 1. M_1 is relatively compact in M with $\partial M_1 = \Sigma_0 \cup C_0$, Σ_0 and C_0 are compact in M with smooth boundary $\partial \Sigma_0 = \partial C_0$. We denote by $H^1_0(\Sigma_0)$ and $H^1_0(C_0)$ the respective restricted Sobolev spaces of order 1, i.e., the spaces of distributions in $H^1(\Sigma_0)$ and $H^1(C_0)$ that vanish on the boundary.

If $f \in H^1_0(\Sigma_0) \oplus L^2(\Sigma_0)$ is a pair of Cauchy data, we denote by $e_{\Sigma_0} f$ its extension by 0 to Σ and by $u = U_{\Sigma_0} f$ the restriction to M_1 of the solution of the Cauchy problem

$$\begin{cases} Pu = 0 & \text{in } M, \\ \rho u = e_{\Sigma_0} f & \text{on } \Sigma, \end{cases}$$

where $\rho u = (u\restriction_\Sigma, \mathrm{i}^{-1}\partial_\nu u\restriction_\Sigma)$. By standard energy estimates one obtains that

$$U_{\Sigma_0} : H^1_0(\Sigma_0) \oplus L^2(\Sigma_0) \to H^1(M_1)$$

is continuous.

In Section 8C we prove the following result.

Theorem 1.2. *The map*

$$T : H^1_0(\Sigma_0) \oplus L^2(\Sigma_0) \to H^1_0(C_0), \qquad f \mapsto (U_{\Sigma_0} f)\restriction_{C_0},$$

is a homeomorphism. Moreover, if $\dim M \geq 4$ *then* $T(C^\infty_0(\Sigma_0) \oplus C^\infty_0(\Sigma_0))$ *is dense in* $|D_s|^{-1/2} L^2(\widetilde{C})$.

The first part of Theorem 1.2 is equivalent to the existence and uniqueness of solutions in M_1 of the characteristic Cauchy problem

$$\begin{cases} Pu = 0 & \text{in } M_1, \\ u\restriction_{C_0} = \varphi, & \varphi \in H_0^1(C_0). \end{cases}$$

The proof proceeds by reduction to a case already considered by Hörmander [1990b], namely when the characteristic surface is the graph of a Lipschitz function defined on a *compact* domain. Beside [Hörmander 1990b] there is a considerable literature on the characteristic Cauchy problem for the Klein–Gordon equation, for example [Bär and Wafo 2015; Cagnac 1981; Dossa 2002; Nicolas 2006]; let us also mention related works on the Dirac equation [Nicolas 2002; Häfner and Nicolas 2011; Joudioux 2011]. The first part of Theorem 1.2 could actually also be deduced from [Bär and Wafo 2015, Theorem 23].

The second part of Theorem 1.2 asserts that there is no loss of information on the level of purity of states when going from the cone C to its interior M_0. The precise form of the statement comes from the fact that the one-particle Hilbert space associated to our Hadamard states, namely, the completion of $\mathcal{H}(\widetilde{C})$ for the inner product $(\cdot \mid (\lambda^+ + \lambda^-)\,\cdot)$, equals $|D_s|^{-1/2}L^2(\widetilde{C})$. The validity of this result appears to be very delicate; it would be for instance problematic for $|D_s|^{-\alpha}L^2(\widetilde{C})$ with $\alpha < \frac{1}{2}$ instead of $\alpha = \frac{1}{2}$ and we do not know whether it holds for $d < 3$. The generalization of Theorem 1.2 to other geometrical situations is thus an interesting open problem, particularly relevant for the quantum field theoretical bulk-to-boundary correspondence.

Plan of the paper. In Section 2 we fix the geometric setup and outline the construction of null coordinates near the cone C. In Section 3 we briefly review the Klein–Gordon field in M_0 and the definition of Hadamard states. Section 4 is devoted to the so-called bulk-to-boundary correspondence, i.e., to the definition of a convenient symplectic space $(\mathcal{H}(\widetilde{C}), \sigma_C)$ of functions on C, containing the traces on C of space-compact solutions in M_0.

In Section 5, we formulate the Hadamard condition on C, that is, the natural microlocal condition on the two-point functions of a quasifree state on $(\mathcal{H}(\widetilde{C}), \sigma_C)$ that ensures that the induced state in M_0 is a Hadamard state.

Section 6 is devoted to the pseudodifferential calculus on $\mathbb{R} \times \mathbb{S}^{d-1}$, more precisely to the "product-type" classes associated to bihomogeneous symbols. We also describe more general operator classes, which are pseudodifferential only in the variables in $\mathbb{S}^{d-1}$.

In Section 7 we construct large classes of Hadamard states on the cone, whose covariances belong to the operator classes introduced in Section 6. In Section 8 we characterize *pure* Hadamard states and show that they induce pure states in M_0. Finally in Section 9 we discuss the invariance of our classes of Hadamard states under change of null coordinates on C. Various technical results are collected in the Appendix.

2. Geometric setup

In this section we describe our geometrical setup and construct null coordinates near the cone C.

2A. ***Future lightcone.*** We consider a globally hyperbolic spacetime (M, g) of dimension $\dim M = d+1$. If $K \subset M$, then $I^{\pm}(K; M)$ and $J^{\pm}(K; M)$ denote the future/past timelike and causal, respectively, shadow of K in M; see, e.g., [Wald 1984, Chapter 8] or [Bär et al. 2007, Section 1.3] for more details. If the spacetime M is clear from the context these sets will simply be denoted by $I^{\pm}(K)$ and $J^{\pm}(K)$.

As outlined in the introduction, we fix a base point $p \in M$ and consider

$$C = \partial J^{+}(p)\backslash\{p\} \quad \text{and} \quad M_0 = I^{+}(p),$$

so that C is the future lightcone from p, with tip removed, and M_0 is the interior of C. From [Wald 1984, Section 8.1] we know that M_0 is open, with

$$\overline{M}_0 = J^{+}(p), \quad \partial M_0 = \partial J^{+}(p) = C \cup \{p\}.$$

We assume Hypothesis 1.1, i.e., that there exists $f \in C^{\infty}(M)$ such that:

(1) $C \subset f^{-1}(\{0\})$, $\nabla_a f \neq 0$ on C, $\nabla_a f(p) = 0$ and $\nabla_a \nabla_b f(p) = -2g_{ab}(p)$.

(2) The vector field $\nabla^a f$ is complete on C.

It follows that C is a smooth hypersurface, although $\overline{C}$ is not smooth. Moreover, since C is a null hypersurface, $\nabla^a f$ is tangent to C.

2B. ***Causal structure.*** We now collect some useful results on the causal structure of M_0 and M.

Lemma 2.1. *Let $K \subset M_0$ be compact. Then*:

$$J^{-}(K) \cap J^{+}(p) \quad \textit{is compact}, \tag{2-1}$$

$$J^{+}(K) \cap C = \varnothing. \tag{2-2}$$

Proof. Equation (2-1) follows from [Bär et al. 2007, Lemma A.5.7]. Moreover, if $V \subset M_0$ is open with $K \subset V$, we have $J^{+}(K) \subset I^{+}(V) \subset M_0$. Since $\partial J^{-}(p) = \partial M_0$ and M_0 is open, this implies (2-2). □

The following lemma is due to Moretti [2006, Theorem 4.1(a)]. If $K \subset M_0$, the notation $J^{\pm}(K; M_0)$ or $J^{\pm}(K; M)$ is used in place of $J^{\pm}(K)$ to specify which causal structure one refers to.

Lemma 2.2. *The Lorentzian manifold (M_0, g) is globally hyperbolic. Moreover,*

$$J^{+}(K; M_0) = J^{+}(K; M) \quad \textit{and} \quad J^{-}(K; M_0) = J^{-}(K; M) \cap M_0 \quad \textit{for all } K \subset M_0. \tag{2-3}$$

The next proposition is also due to Moretti [2008, Lemma 4.3].

Proposition 2.3. *Let $K \subset M_0$ be compact. Then there exists a neighborhood U_1 of p in M such that no null geodesic starting from K intersects $\overline{C} \cap U_1$.*

2C. ***Asymptotically flat spacetimes.*** In what follows we explain the relation between Hypothesis 1.1 and the geometrical assumptions met in the literature on Hadamard states [Moretti 2006; 2008; Dappiaggi and Siemssen 2013; Benini et al. 2014].

Let us consider two globally hyperbolic spacetimes (M_0, g_0) and (M, g), where M_0 is an embedded submanifold of M. One introduces the following set of assumptions:

Hypothesis 2.4. Suppose the spacetime (M, g) is such that

(1) there exists $\Omega \in C^\infty(M)$ with $\Omega > 0$ on M_0 and $g\restriction_{M_0} = \Omega^2\restriction_{M_0} g_0$,

(2) there exists $i^- \in M$ such that $J^+(i^-; M)$ is closed and

$$M_0 = J^+(i^-; M)\backslash \partial J^+(i^-; M),$$

(3) g_0 solves the vacuum Einstein equations at least in a neighborhood of

$$\mathscr{I}^- := \partial J^+(i^-; M)\backslash\{i^-\},$$

(4) $\Omega = 0$ and $d\Omega \neq 0$ on $\mathscr{I}^-$, $d\Omega(i^-) = 0$ and $\nabla_a\nabla_b\Omega(i^-) = -2g_{ab}(i^-)$,

(5) if $n^a := g^{ab}\nabla_b\Omega$, then there exists $\omega \in C^\infty(M)$ with $\omega > 0$ on $M_0 \cup \mathscr{I}^-$ and

(a) $\nabla_a(\omega^4 n^a) = 0$ on $\mathscr{I}^-$,

(b) the vector field $\omega^{-1}n$ is complete on $\mathscr{I}^-$.

Above, the symbols ∇_a refer to the metric g.

One says that (M_0, g_0) is an *asymptotically flat vacuum spacetime with past time infinity* i^- if there exists a spacetime (M, g) such that M_0 is an embedded submanifold of M and Hypothesis 2.4 is satisfied.[3]

Lemma 2.5. *Suppose* (M_0, g_0) *is an asymptotically flat vacuum spacetime with past time infinity* i^- *and let* (M, g) *satisfy Hypothesis 2.4. Then Hypothesis 1.1 is satisfied for* $p := i^-$ *and* $f = \omega\Omega$.

Note that actually only conditions (1), (2), (4) and (5b) in Hypothesis 2.4 are needed in Lemma 2.5.

In the present paper we construct Hadamard states for the Klein–Gordon operator $P = -\Box_g + r(x)$ in $(M_0, g\restriction_{M_0})$ for any smooth, real-valued r. In the special case of the conformal wave operator $P = -\Box_g + (n-2)/(4(n-1))\,R$ (with R the scalar curvature) this yields, however, also Hadamard states on (M_0, g_0), since the two metrics are conformally related; see Appendix A2.

2D. ***Null coordinates near C.*** For later use it is convenient to introduce null coordinates near C. The construction seems to be well known; we sketch it for the reader's convenience. Note however the estimates in Lemma 2.6, which will be useful later on.

We first choose normal coordinates $(y^0, \bar{y})$ at p such that $C = \{(y^0, \bar{y}) \mid (y^0)^2 - |\bar{y}|^2 = 0,\ y^0 > 0\}$ on a neighborhood of p.

Set

$$v := y^0 + |\bar{y}|, \quad w := y^0 - |\bar{y}|, \quad \psi := \frac{\bar{y}}{|\bar{y}|} \in \mathbb{S}^{d-1}, \tag{2-4}$$

[3]Note that we consider here only globally hyperbolic spacetimes; see [Moretti 2008, Appendix A] for a more general definition.

so that on a neighborhood of p one has $C = \{w = 0,\ v > 0\}$. Abusing notation slightly, we denote by $\psi^1, \dots, \psi^{d-1}$ coordinates on $\mathbb{S}^{d-1}$ and use the same letter for their pullback to local coordinates on M near p. We set

$$S := \{w = 0,\ v = \epsilon_0\}, \tag{2-5}$$

where $\epsilon_0 > 0$ will be chosen to be small enough. Note that $S \subset C$ is diffeomorphic to $\mathbb{S}^{d-1}$.

Lemma 2.6. (1) *There exists a unique solution $s \in C^\infty(C)$ of*

$$\begin{cases} (\nabla^a f \nabla_a s)\restriction_C = -1, \\ s\restriction_S = 0. \end{cases}$$

(2) *There exists unique solutions $\theta^j \in C^\infty(C)$, $1 \le j \le d-1$, of*

$$\begin{cases} (\nabla^a f \nabla_a \theta^j)\restriction_C = 0, \\ \theta^j\restriction_S = \psi^j. \end{cases}$$

(3) *Moreover, there exists $0 < \epsilon_0 < \epsilon_1$ and $k, \tilde{\theta}^j \in C^\infty(]-\epsilon_1, \epsilon_1[\, \times \mathbb{S}^{d-1})$ such that*

$$s(v, \psi) = \tfrac{1}{2}\ln(v) + k(v, \psi) \quad \textit{and} \quad \theta^j(v, \psi) = \tilde{\theta}^j(v, \psi) \quad \textit{on }]0, \epsilon_0[\, \times \mathbb{S}^{d-1}.$$

Proof. The proof is given in Appendix A4. □

It remains to extend s and θ^j to smooth functions on a neighborhood of C.

We argue as in [Wald 1984, Section 11.1]: for $s_0 \in \mathbb{R}$, the submanifold $S_{s_0} = \{s = s_0\} \subset C$ is spacelike, of codimension 2 in M. At a given point of S_{s_0} the orthogonal to its tangent space is two-dimensional and timelike, and hence contains two null lines. One of them is generated by $\nabla^a f$; the other is transverse to C. We extend (s, θ) to a neighborhood of C by imposing that (s, θ) are constant along the above family of null geodesics, transverse to C.

Lemma 2.7. *The functions (f, s, θ) constructed above are a system of local coordinates near C with $C \subset \{f = 0\}$ and*

$$g\restriction_C = -2\, df\, ds + h_{ij}(s, \theta)\, d\theta^i\, d\theta^j, \tag{2-6}$$

where $h_{ij}(s, \theta)\, d\theta^i\, d\theta^j$ is a smooth, s-dependent Riemannian metric on $\mathbb{S}^{d-1}$.

Proof. The proof will be given in Appendix A3. □

2E. *Estimates on traces.* In this subsection we derive estimates, in the coordinates (s, θ) on C constructed above, for the restriction to C of a smooth, space-compact function in M. These estimates will be applied later to traces on C of solutions of the Klein–Gordon equation in M_0.

We recall that $C^\infty_{\mathrm{sc}}(M)$ denotes the space of smooth *space compact* functions, i.e., the space of $\phi \in C^\infty(M)$ such that $\operatorname{supp} \phi \subset J^+(K) \cup J^-(K)$ for some compact $K \subset M$.

We will slightly abuse notation by writing $\phi(x^0, \dots, x^d)$ for the function ϕ expressed in some coordinate system $(x^0, \dots, x^d)$ near p. We will similarly write, for example, $\phi(v, \psi)$ or $\phi(s, \theta)$ for $\phi \in C^\infty(C)$.

By Lemma 2.1 we see that $\operatorname{supp} \phi \cap \overline{C}$ is compact in $\overline{C}$ if $\phi \in C^\infty_{\mathrm{sc}}(M)$. This means that it suffices to control the derivatives in (s, θ) of $\phi\restriction_C\ (s, \theta)$ near $s = -\infty$, that is, of $\phi\restriction_C\ (v, \psi)$ near $v = 0$. Clearly the only task is to control what happens near p, that is, when $s \to -\infty$. We first derive estimates in the coordinates (v, ψ) introduced in (2-4) in a neighborhood of $v = 0$. If $\phi \in C^\infty_{\rm sc}(M)$ we denote by $\phi(y^0, \bar{y})$ the function ϕ expressed in normal coordinates at p, which is defined on a neighborhood of 0. We then set

$$\hat{\phi}(v, \psi) = \phi\big(\tfrac{1}{2}v, \tfrac{1}{2}v\psi\big) \in C^\infty(]-\epsilon_1, \epsilon_1[\times \mathbb{S}^{d-1}) \quad \text{for some } \epsilon_1 > 0,$$

so that

$$\phi\restriction_C = \hat{\phi}\restriction_{\{v>0\}}.$$

We denote by S^0 the space of functions $u(v, \psi) \in C^\infty(]-\epsilon_1, \epsilon_1[\times \mathbb{S}^{d-1})$ which are bounded with all derivatives.

Lemma 2.8. (1) *If $\phi \in C^\infty_{\rm sc}(M)$ then $\hat{\phi}(v, \psi)$ belongs to S^0.*

(2) *Let $|h| = \det[h_{ij}]$. Then $|h|(v, \psi) = v^{2(d-1)} r_0(v, \psi)$ for $r_0, r_0^{-1} \in S^0$.*

Proof. Considering the map $\chi : \mathbb{S}^{d-1} \to \mathbb{R}^d$, $\psi \mapsto \psi$, and still denoting by ψ some coordinates, on $\mathbb{S}^{d-1}$ we have

$$\partial_v \tilde{\phi} = \tfrac{1}{2}(\partial_{y^0}\phi - \psi \cdot \partial_{\bar{y}}\phi) \quad \text{and} \quad \partial_{\psi^i}\tilde{\phi} = \tfrac{1}{2} v \partial_{\psi^i}\chi^j \partial_{\bar{y}^j}\phi.$$

From this we obtain (1). To prove (2) we need to express $h_{ij} = \langle \partial_{\theta^i} \mid g\, \partial_{\theta^j}\rangle$ on C. An easy computation using the estimates in Lemma 2.6 shows that on C we have

$$\partial_{\theta^i} = a_i^j(v, \psi)\,\partial_{\psi^j} + v r_0(v, \psi)\,\partial_v,$$

where $a_i^j, r_0 \in S^0$ and $[a^{ij}](v, \psi)$ is invertible. Plugging this into (A-9), we obtain

$$[h_{ij}](v, \psi) = v^2\big({}^t[a_i^j](v, \psi)[m_{ij}](\psi)[a_i^j](v, \psi) + v[b_{ij}](v, \psi)\big),$$

where $b_{ij} \in S^0$. This implies (2). □

Later we will also need the following lemma. We denote by $m_{ij}(\theta)\, d\theta^i\, d\theta^j$ the standard Riemannian metric on $\mathbb{S}^{d-1}$ and set

$$\beta(s, \theta) := |m|^{1/4}(\theta)|h|^{-1/4}(s, \theta). \tag{2-7}$$

Lemma 2.9. *Let*

$$\tilde{\phi}(s, \theta) := \beta^{-1}(s, \theta)\phi\restriction_C(s, \theta), \quad \phi \in C^\infty_{\rm sc}(M).$$

Then for all $s_1 \in \mathbb{R}$ one has

$$\partial_s^\alpha \partial_\theta^\beta \tilde{\phi} \in O(\mathrm{e}^{s(d-1)}), \quad s \in]-\infty, s_1], \quad \textit{for all } \alpha, \beta.$$

Proof. We note that $\beta^{-1} = v^{(d-1)/2} r_0(v, \psi)$, for $r_0, r_0^{-1} \in S^0$. From this and Lemma 2.8, it follows that if $\phi \in C_0^\infty(M)$ then $\tilde{\phi}(v, \psi) \in v^{(d-1)/2} S^0$. It remains to estimate the derivatives of $\tilde{\phi}$ with respect to s and θ. By a standard computation we obtain, for $u \in C^\infty(]-\epsilon_1, \epsilon_1[\times \mathbb{S}^{d-1})$,

$$\partial_{\theta^i} u = a_i^j(v, \psi)\,\partial_{\psi^j} u + v r_i(v, \psi)\,\partial_v u, \quad \text{and} \quad \partial_s u = v(1 + v r_0(v, \psi))\,\partial_v u + v b^j(v, \psi)\,\partial_{\psi^j} u$$

for $r_0, r_i, b^j, a_i^j \in S^0$ and $[a_i^j]$ invertible. From this point on the lemma is a routine computation. □

3. Klein–Gordon fields inside the future lightcone

3A. ***Klein–Gordon equation in M_0.*** We fix a smooth real function $r \in C^\infty(M)$ and consider the *Klein–Gordon operator* on (M, g)

$$P(x, D_x) = -\nabla^a \nabla_a + r(x) \quad \text{acting on } C^\infty(M).$$

We denote by $E_\pm$ in $\mathscr{D}'(M \times M)$ the retarded and advanced Green's functions for P, by $E = E_+ - E_-$ in $\mathscr{D}'(M \times M)$ the Pauli–Jordan commutator function and by $\mathrm{Sol}_{\mathrm{sc}}(P)$ the space of smooth, complex-valued, space-compact solutions of

$$P(x, D_x)\phi = 0 \quad \text{in } M.$$

Recall that we have set in Section 2A

$$M_0 := I^+(p)$$

and, by Lemma 2.2, we know that (M_0, g) is globally hyperbolic.

We denote by $P_0 = -\nabla^a \nabla_a + r(x)$ the restriction of P to M_0, by $E_0 \in \mathscr{D}'(M_0 \times M_0)$ the Pauli–Jordan function for P_0 and by $\mathrm{Sol}_{\mathrm{sc}}(P_0)$ the space of smooth, complex-valued, space-compact solutions of

$$P_0(x, D_x)\phi_0 = 0 \quad \text{in } M_0.$$

By the global hyperbolicity of (M_0, g) we know that $\mathrm{Sol}_{\mathrm{sc}}(P_0) = E_0\mathscr{D}(M_0)$. From (2-3) and the uniqueness of $E_{0\pm}$ we obtain that $E_{0\pm} = E_\pm\restriction_{M_0 \times M_0}$; hence,

$$E_0 = E\restriction_{M_0 \times M_0} .$$

It follows that any $\phi_0 \in \mathrm{Sol}_{\mathrm{sc}}(P_0)$ uniquely extends to $\phi \in \mathrm{Sol}_{\mathrm{sc}}(P)$; in fact,

$$\phi_0 = E_0 f_0, \;\; f_0 \in \mathscr{D}(M_0) \quad \Longrightarrow \quad \phi_0 = E f_0\restriction_{M_0} . \tag{3-1}$$

As usual we equip $\mathrm{Sol}_{\mathrm{sc}}(P_0)$ with the symplectic form

$$\bar{\phi}_1 \sigma_0 \phi_2 := \int_{\Sigma_0} \overline{\nabla_a \phi_1} \phi_2 - \bar{\phi}_1 \nabla_a \phi_2 n^a \, d\sigma_h, \tag{3-2}$$

where $\Sigma_0 \subset M_0$ is a Cauchy hypersurface for (M_0, g) (see Appendix A1 for notation). It is well known that

$$E_0 : (C_0^\infty(M_0)/P_0 C_0^\infty(M_0), E_0) \to (\mathrm{Sol}_{\mathrm{sc}}(P_0), \sigma_0)$$

is a symplectomorphism.

3B. ***Hadamard states in M_0.*** We first briefly recall some standard facts and refer, for example, to [Gérard and Wrochna 2014, Section 2] for details and notation.

3B1. *Covariances of a quasifree state.* If $(\mathcal{Y}, \sigma)$ is a complex symplectic space, the *complex covariances* $\Lambda^{\pm} \in L_{\rm h}(\mathcal{Y}, \mathcal{Y}^*)$ of a (gauge-invariant) quasifree state ω on $\mathrm{CCR}(\mathcal{Y}, \sigma)$ (the polynomial CCR *-algebra of $(\mathcal{Y}, \sigma)$) are defined by

$$\omega(\psi(y_1)\psi^*(y_2)) =: (y_1 \mid \Lambda^+ y_2), \quad \omega(\psi^*(y_2)\psi(y_1)) =: (y_1 \mid \Lambda^- y_2), \quad y_1, y_2 \in \mathcal{Y}.$$

From the CCR we obtain that $\Lambda^+ - \Lambda^- = \mathrm{i}\sigma =: q$, and the necessary and sufficient condition for $\Lambda^{\pm}$ to be the complex covariances of a (gauge-invariant) quasifree state is that $\Lambda^{\pm} \geq 0$.

If $(\mathcal{Y}, \sigma) = (C_0^\infty(M_0)/PC_0^\infty(M_0), E_0)$, the complex covariances of a state ω are induced from *two-point functions*, still denoted by $\Lambda^{\pm}$, such that

$$\Lambda^{\pm} \in \mathcal{D}'(M_0 \times M_0), \quad P\Lambda^{\pm} = \Lambda^{\pm} P = 0,$$

where we identify operators on $C_0^\infty(M_0)$ with sesquilinear forms using the scalar product

$$(u \mid v) := \int_{M_0} \bar{u} v \, d\mu_g, \quad u, v \in C_0^\infty(M_0).$$

3B2. *Hadamard condition.* We now recall the *Hadamard condition* for quasifree states. We denote by T^*M the cotangent bundle of M and by $Z = \{(x, 0)\} \subset T^*M$ the zero section. The *principal symbol* of P is $p(x, \xi) = \xi_a g^{ab}(x) \xi_b$; the set

$$\mathcal{N} := \{(x, \xi) \in T^*M \backslash Z : p(x, \xi) = 0\}$$

is called the *characteristic manifold* of p.

The Hamilton vector field of p will be denoted by H_p, whose integral curves inside $\mathcal{N}$ are called *bicharacteristics*.

We will use the notation $X = (x, \xi)$ for points in $T^*M \backslash Z$ and write $X_1 \sim X_2$ if $X_1 = (x_1, \xi_1)$ and $X_2 = (x_2, \xi_2)$ are in $\mathcal{N}$ and X_1 and X_2 lie on the same bicharacteristic of p.

Let us fix a time orientation and denote by $V_{x\pm} \subset T_x M$ for $x \in M$ the open future/past lightcones and $V^*_{x\pm}$ the dual cones

$$V_x^{*\pm} := \{\xi \in T_x^* M : \xi \cdot v > 0 \text{ for all } v \in V_{x\pm} \text{ with } v \neq 0\}.$$

The set $\mathcal{N}$ has two connected components invariant under the Hamiltonian flow of p, namely

$$\mathcal{N}^{\pm} := \{X \in \mathcal{N} : \xi \in V_x^{*\pm}\}.$$

Definition 3.1. A quasifree state ω on $\mathrm{CCR}(C_0^\infty(M_0)/PC_0^\infty(M_0), E_0)$ with two-point functions $\Lambda^{\pm}$ satisfies the *microlocal spectrum condition* if

$$\mathrm{WF}(\Lambda^{\pm})' \subset \mathcal{N}^{\pm} \times \mathcal{N}^{\pm}. \tag{μsc}$$

Quasifree states satisfying (μsc) are called *Hadamard states.*

This form of the Hadamard condition was shown in [Sahlmann and Verch 2001] to be equivalent to older definitions [Radzikowski 1996]; we refer the reader to [Sanders 2010; Wrochna 2013] for a discussion on equivalent formulations of the microlocal spectrum condition.

4. Bulk-to-boundary correspondence

4A. ***Boundary symplectic space.*** We equip C with the coordinates (s, θ) constructed in Section 2D and hence identify C with

$$\widetilde{C} := \mathbb{R} \times \mathbb{S}^{d-1}. \tag{4-1}$$

We denote by $H^k(\widetilde{C})$, $k \in \mathbb{N}$, the Sobolev space

$$H^k(\widetilde{C}) := \left\{ g \in \mathscr{D}'(\mathbb{R} \times \mathbb{S}^{d-1}) : \int |\partial_s^\alpha \partial_\theta^\beta g|^2 |m|^{1/2}\, ds\, d\theta < \infty,\ \alpha + |\beta| \le k \right\},$$

and extend the definition of $H^k(\widetilde{C})$ to $k \in \mathbb{R}$ in the usual way. The space $H^0(\widetilde{C})$ will be denoted simply by $L^2(\widetilde{C})$. We set also

$$\mathscr{H}(\widetilde{C}) := \bigcap_{k\in\mathbb{R}} H^k(\widetilde{C}) \quad \text{and} \quad \mathscr{H}'(\widetilde{C}) := \bigcup_{k\in\mathbb{R}} H^k(\widetilde{C}),$$

equipped with their canonical topologies.

We set

$$\bar{g}_1 \sigma_C g_2 := \int_{\mathbb{R}\times\mathbb{S}^{d-1}} (\partial_s \bar{g}_1 g_2 - \bar{g}_1 \partial_s g_2)|m|^{1/2}(\theta)\, ds\, d\theta, \quad g_1, g_2 \in \mathscr{H}(\widetilde{C}). \tag{4-2}$$

Introducing the *charge* $q := \mathrm{i}\sigma_C$ we have

$$\bar{g}_1 q g_2 = 2(g_1 \mid D_s g_2)_{L^2(\widetilde{C})}, \quad g_1, g_2 \in \mathscr{H}(\widetilde{C}),$$

where $D_s = \mathrm{i}^{-1}\partial_s$ is selfadjoint on $L^2(\widetilde{C})$ on its natural domain. Clearly $(\mathscr{H}(\widetilde{C}), \sigma_C)$ is a complex symplectic space.

4B. ***Bulk-to-boundary correspondence.***

Definition 4.1. Let $\beta \in C^\infty(\widetilde{C})$ be as defined in (2-7). We set

$$\rho : \mathrm{Sol}_{\mathrm{sc}}(P_0) \to C^\infty(\mathbb{R} \times \mathbb{S}^{d-1}), \quad \phi \mapsto \beta^{-1}(s, \theta)\phi\restriction_C(s, \theta).$$

Proposition 4.2. (1) *ρ maps $\mathrm{Sol}_{\mathrm{sc}}(P_0)$ into $\mathscr{H}(\widetilde{C})$.*

(2) *$\rho : (\mathrm{Sol}_{\mathrm{sc}}(P_0), \sigma) \to (\mathscr{H}(\widetilde{C}), \sigma_C)$ is a monomorphism, i.e.,*

$$\overline{\rho\phi}_1 \sigma_C \rho\phi_2 = \bar{\phi}_1 \sigma \phi_2 \quad \textit{for all } \phi_1, \phi_2 \in \mathrm{Sol}_{\mathrm{sc}}(P_0).$$

Proof. Let ϕ_0 and ϕ be as in (3-1). By Lemma 2.1 and the support properties of E, we see that $\operatorname{supp}\phi \cap \overline{C}$ is compact in M. Therefore the restriction of ϕ to C equals the restriction of a smooth, compactly supported function to C. By Lemma 2.9 and the fact that $\rho\phi_0$ is supported in $]-\infty, s_1[\, \times \mathbb{S}^{d-1}$ for some s_1, we obtain that $\rho\phi_0 \in \mathscr{H}(\widetilde{C})$, which proves (1).

We now prove (2). Let $\phi_{i,0} \in \mathrm{Sol}_{\mathrm{sc}}(P_0)$, $i = 1, 2$, be restrictions to M_0 of $\phi_i \in \mathrm{Sol}_{\mathrm{sc}}(P)$. We fix a Cauchy surface Σ_0 for (M_0, g) such that $\operatorname{supp}\phi_{i,0} \cap \Sigma_0 \subset K \Subset M_0$. We can find a Cauchy surface Σ for (M, g) such that $\Sigma \cap K = \Sigma_0 \cap K$. Denoting by

$$J_a(\phi_1, \phi_2) := \bar{\phi}_1 \nabla_a \phi_2 - \overline{\nabla_a \phi_1}\phi_2,$$

the conserved current, we have

$$\bar{\phi}_{1,0}\sigma_0\phi_{2,0} = \bar{\phi}_1\sigma\phi_2,$$

where

$$\bar{\phi}_1\sigma\phi_2 = -\int_\Sigma J_a(\phi_1, \phi_2)n^a\, d\sigma_h$$

is the symplectic form on $\mathrm{Sol}_{\mathrm{sc}}(P)$. We now apply Stokes formula in the form (A-6) to the domain $U \subset M$ bounded by $\Sigma \cap K$, $\overline{C}$ and $\partial J^+(\Sigma \cap K)$, using that $\nabla_a J^a(\phi_1, \phi_2) = 0$. The boundary term on $\Sigma \cap K$ yields $-\bar{\phi}_1\sigma\phi_2$; the boundary term on $\partial J^+(\Sigma \cap K)$ vanishes. To express the boundary term on $\overline{C}$, we use the coordinates (f, s, θ) constructed in Section 2D. We formally obtain the quantity

$$\bar{g}_1\hat{\sigma}g_2 = \int_{\mathbb{R}\times\mathbb{S}^{d-1}} (\partial_s\bar{g}_1 g_2 - \bar{g}_1\,\partial_s g_2)|h|^{1/2}(s, \theta)\, ds\, d\theta$$

for $g_i = (\phi_i)\restriction_C$. This equals $\overline{\rho\phi}_1\sigma_C\rho\phi_2$ by an easy computation.

To justify the use of Stokes formula, we need to take care of the fact that $\overline{C}$ is not smooth at p. This can be done as follows: for $0 < \epsilon \ll 1$, we denote by U_ϵ some ϵ-neighborhood of p. We replace $\overline{C}$ by a smooth hypersurface C_ϵ, obtained by smoothly gluing $C\backslash U_\epsilon$ to a piece of a Cauchy surface Σ'_ϵ passing through U_ϵ. The contribution of the integral on Σ_ϵ is written using (A-4) and converges to 0 when $\epsilon \to 0$, using that ϕ_i are smooth functions. The contribution of the integral on $C\backslash U_\epsilon$ converges to $\overline{\rho\phi}_1\sigma_C\rho\phi_2$, using that $\rho\phi_i \in \mathscr{H}(\widetilde{C})$. This completes the proof of the proposition. ☐

4C. *Pullback of states from the boundary.* Since

$$\rho : (\mathrm{Sol}_{\mathrm{sc}}(P_0), \sigma_0) \to (\mathscr{H}(\widetilde{C}), \sigma_C)$$

is a monomorphism, we can pull back a quasifree state ω_C on $\mathrm{CCR}(\mathscr{H}(\widetilde{C}), \sigma_C)$ to a quasifree state ω_0 on $\mathrm{CCR}(C_0^\infty(M_0)/P_0C_0^\infty(M_0), E_0)$ by setting

$$\omega_0(\psi(u_1)\psi^*(u_2)) := \omega_C\,(\psi(\rho\circ E_0u_1)\psi^*(\rho\circ E_0u_2)), \quad u_1, u_2 \in C_0^\infty(M_0). \tag{4-3}$$

If $\lambda^\pm \in L_{\mathrm{h}}(\mathscr{H}(\widetilde{C}), \mathscr{H}(\widetilde{C})^*)$ are the complex covariances of ω_C, then the complex covariances of ω_0 are (formally) given by

$$\Lambda^\pm := (\rho\circ E_0)^*\circ\lambda^\pm\circ(\rho\circ E_0). \tag{4-4}$$

5. Hadamard condition on the cone

In this section we formulate the natural boundary version of the bulk Hadamard condition (μsc).

5A. *Preparations.* We recall that $p(x, \xi)$ denotes the principal symbol of the Klein–Gordon operator P (or P_0).

Let $C \subset M$ be the forward lightcone introduced in Section 2A. We denote by $N^*C \subset T^*M\backslash Z$ the *conormal bundle* to C, namely,

$$N^*C := \{(x, \xi) \in T^*M\backslash Z : x \in C \text{ and } \xi = 0 \text{ on } T_xC\}.$$

The fact that C is characteristic is equivalent to

$$N^*C \subset \mathcal{N}, \tag{5-1}$$

where $\mathcal{N}$ is the characteristic manifold of p. Since N^*C is Lagrangian, it is well known that (5-1) implies that N^*C is invariant under the flow of H_p. The projections on M of bicharacteristics starting from N^*C are (modulo reparametrization) *characteristic curves*, i.e., integral curves of the vector field $v^a = \nabla^a f$ if $f \in C^\infty(M)$ is some defining function of C, that is, $f = 0$ and $df \neq 0$ on C.

We will use the coordinates (f, s, θ) introduced in Section 2D, which, for ease of notation, will be denoted by $x = (r, s, y) \in \mathbb{R} \times \mathbb{R} \times \mathbb{S}^{d-1}$. The dual coordinates are denoted by $\xi = (\varrho, \sigma, \eta)$, elements of T^*M will sometimes be denoted by $X = (x, \xi)$ and elements of T^*C will be denoted by $Y = ((s, y), (\sigma, \eta))$.

In the above coordinates, we have

$$C = \{r = 0\} \quad \text{and} \quad N^*C = \{r = 0,\ \sigma = \eta = 0\}$$

and, from (2-6), we obtain that

$$p(x, \xi)\restriction_C = -2\varrho\sigma + h(s, y, \eta), \tag{5-2}$$

where we set $h(s, y, \eta) = h^{ij}(0, s, y)\eta_i\eta_j$. Note that $h(s, y, \eta)$ is *elliptic*, that is, $h(s, y, \eta) \geq c_0|\eta|^2$ for $c_0 > 0$, locally in (s, y), since $h_{ij}\, dy^i\, dy^j$ is Riemannian.

For later use let us extend the notation $X_1 \sim X_2$ introduced in Section 3B2. For $Y = (s, y, \sigma, \eta) \in T^*C$ and $X = (x, \xi) \in T^*M$, we will write $Y \sim X$ if

$$\sigma \neq 0 \quad \text{and} \quad ((0, s, y), ((2\sigma)^{-1}h(s, y, \eta), \sigma, \eta)) \sim X. \tag{5-3}$$

Recall also that the positive/negative energy components $\mathcal{N}^\pm$ of $\mathcal{N}$ were defined in Section 3B2.

Lemma 5.1. *Let $Y_1 = (s_1, y_1, \sigma_1, \eta_1) \in T^*C$ and $X_2 = (x_2, \xi_2) \in T^*M$ with $x_2 \notin C$. Then*:

(1) *There exists $\varrho_1 \in \mathbb{R}$ such that*

$$X_1 := ((0, s_1, y_1), (\varrho_1, \sigma_1, \eta_1)) \sim (x_2, \xi_2) =: X_2$$

if and only if $\sigma_1 \neq 0$, in which case $\varrho_1 = (2\sigma_1)^{-1}h(s_1, y_1, \eta_1)$ and $Y_1 \sim X_2$.

(2) *If $Y_1 \sim X_2$, then $X_2 \in \mathcal{N}^\pm$ if and only if $\pm\sigma_1 > 0$.*

Proof. Let $X_1 = ((0, s_1, y_1), (\varrho_1, \sigma_1, \eta_1)) \in \mathcal{N}$. By (5-2) we have

$$-2\varrho_1\sigma_1 + h(s_1, y_1, \eta_1) = 0.$$

If $\sigma_1 = 0$ then $h(s_1, y_1, \eta_1) = 0$, hence $\eta_1 = 0$ by ellipticity of h. Therefore $\sigma_1 = 0$ implies $X_1 \in N^*C$. Since $X_2 \sim X_1$ and N^*C is invariant under the flow of H_p, we also have $X_2 \in N^*C$, which contradicts the hypothesis that $x_2 \notin C$. Therefore, necessarily $\sigma_1 \neq 0$, and hence $\varrho_1 = (2\sigma_1)^{-1}h(s_1, y_1, \eta_1)$ and $Y_1 \sim X_2$. This proves (1).

To prove (2) we have to show that

$$\pm\sigma_1 > 0 \iff ((0, s_1, y_1), ((2\sigma_1)^{-1} h(s_1, y_1, \eta_1), \sigma_1, \eta_1)) \in \mathcal{N}^{\pm}. \tag{5-4}$$

Let us fix $(y_1, \eta_1) \in T^*\mathbb{S}^{d-1}$ and $\sigma_1 \in \mathbb{R}$. Since $\mathcal{N}^{\pm}$ are the two connected components of $\mathcal{N}$, it suffices, by connectivity, to prove (5-4) for s_1 in a neighborhood of $-\infty$, i.e., in a neighborhood of p in M. Recall that we introduced Gaussian normal coordinates $(y^0, \bar{y})$ near p with ∂_{y^0} future oriented. Let α be the one-form $(2\sigma_1)^{-1} h(s_1, y_1, \eta_1)\, dr + \sigma_1\, ds + \eta_1\, dy$. Then

$$((0, s_1, y_1), ((2\sigma_1)^{-1} h(s_1, y_1, \eta_1), \sigma_1, \eta_1)) \in \mathcal{N}^{\pm} \iff \mp\langle \alpha \mid g^{-1}\, dy^0\rangle > 0.$$

Since it suffices to check the sign of $\langle \alpha \mid g^{-1}\, dy^0\rangle$ near p, we can, by a simple approximation argument (see, e.g., (A-9)) replace g by the flat metric at p. We then have — see Lemma 2.6 and recall that $s = u$ and $r = f$ —

$$y^0 = v + w, \quad v = \mathrm{e}^{s}, \quad w = \mathrm{e}^{-s} r,$$

hence

$$\mp\langle \alpha \mid g^{-1}\, dy^0\rangle = \pm 2(\mathrm{e}^{-s_1}\sigma_1 + \mathrm{e}^{s_1}(2\sigma_1)^{-1} h(s_1, y_1, \eta_1))$$

has the same sign as $\pm\sigma_1$, which proves (5-4). □

Recall that $E \in \mathcal{D}'(M \times M)$ is the Pauli–Jordan commutator function for P and $\rho : \mathcal{D}(M) \to C^\infty(\widetilde{C})$, $u \mapsto u\restriction_C$, is (modulo a smooth, nonzero multiplicative factor) the operator of restriction to C, defined in Definition 4.1.

Let us recall some notation: identifying $T^*(M_1 \times M_2)$ with $T^*M_1 \times T^*M_2$, we write $(T^*M_1 \times T^*M_2)\backslash Z$ for the image of $T^*(M_1 \times M_2)\backslash Z$ under this identification. If $\Gamma \subset (T^*M_1 \times T^*M_2)\backslash Z$, one sets

$$\begin{aligned} {}_{M_1}\Gamma &:= \{(x_1, \xi_1) : (x_1, \xi_1, x_2, 0) \in \Gamma \text{ for some } x_2\} \subset T^*M_1\backslash Z_1, \\ \Gamma_{M_2} &:= \{(x_2, \xi_2) : (x_1, 0, x_2, \xi_2) \in \Gamma \text{ for some } x_1\} \subset T^*M_2\backslash Z_2, \end{aligned} \tag{5-5}$$

where Z_i is the zero section of T^*M_i.

Proposition 5.2. *Let* $\chi \in C_0^\infty(M)$ *with* $\operatorname{supp}\chi \subset M\backslash C$ *and* $\psi \in C_0^\infty(\widetilde{C})$. *Then*:

(1) $\mathrm{WF}(\psi\rho \circ E\chi)' \subset \{(Y_1, X_2) : y_1 \in \operatorname{supp}\psi,\ x_2 \in \operatorname{supp}\chi,\ Y_1 \sim X_2\}$, *where the notation* $Y \sim X$ *is as defined in* (5-3).

(2) $\psi\rho \circ E\chi : \mathcal{D}(M) \to \mathcal{D}(\widetilde{C})$ *extends continuously as* $\psi\rho \circ E\chi : \mathcal{D}'(M) \to \mathcal{D}'(\widetilde{C})$.

Proof. It is well known that

$$\begin{aligned} \operatorname{supp} E &\subset \{(x_1, x_2) : x_1 \in J(x_2)\}, \\ \mathrm{WF}(E)' &= \{(X_1, X_2) \in \mathcal{N} \times \mathcal{N} :\ X_1 \sim X_2\}. \end{aligned} \tag{5-6}$$

On the other hand, the distributional kernel of ρ equals

$$\delta(r_2) \otimes \delta(s_1, y_1, s_2, y_2)\beta^{-1}(s_1, y_1) \in \mathcal{D}'(\widetilde{C} \times M).$$

It follows that

$$\mathrm{WF}(\rho)' = \{(Y_1, X_2) : r_2 = 0,\ (s_1, y_1) = (s_2, y_2),\ (\sigma_1, \eta_1) = (\sigma_2, \eta_2),\ (\sigma_2, \eta_2) \neq (0,0)\}. \tag{5-7}$$

Since $E : \mathscr{D}(M) \to \mathscr{E}(M)$, we see that $\psi\rho \circ E\chi : \mathscr{D}(M) \to \mathscr{D}(\widetilde{C})$. Moreover, there exists $\chi_1 \in C_0^\infty(M)$ such that $\psi\rho \circ E\chi = \psi\rho \circ \chi_1 E\chi$. We then have

$${}_{\widetilde{C}}\mathrm{WF}(\rho)' = \mathrm{WF}(E)'_M = \varnothing$$

and it follows from [Hörmander 1990a, Chapter 8] and (5-6)–(5-7) that

$$\begin{aligned}\mathrm{WF}(\psi\rho \circ E\chi)' &\subset \mathrm{WF}(\psi\rho)' \circ \mathrm{WF}(E\chi)' \\ &\subset \{(Y_1, X_2) : ((0, s_1, y_1), (\varrho_1, \sigma_1, \eta_1)) \sim X_2 \text{ for some } \varrho_1,\ x_2 \in \operatorname{supp}\chi\}.\end{aligned}$$

Using that $\operatorname{supp}\chi \cap C = \varnothing$ and Lemma 5.1(1), this implies (1). Moreover, (1) implies that

$$\mathrm{WF}(\psi\rho \circ E\chi)'_M = \varnothing. \tag{5-8}$$

Again by [Hörmander 1990a], this implies that $\psi\rho \circ E\chi = \mathscr{D}(M) \to \mathscr{D}(\widetilde{C})$ extends continuously as $\psi\rho \circ E\chi : \mathscr{D}'(M) \to \mathscr{D}'(\widetilde{C})$. □

5B. ***Hadamard condition on the cone.*** Recall from Section 4B that we can associate to a quasifree state ω_C on $\mathrm{CCR}(\mathscr{H}(\widetilde{C}), \sigma_C)$ a quasifree state ω_0 on $\mathrm{CCR}(C_0^\infty(M_0)/PC_0^\infty(M_0), E_0)$. In this subsection we give natural conditions on the covariances $\lambda^\pm$ of ω_C which ensure that the induced state ω_0 satisfies the microlocal spectrum condition (μsc).

Recall that we denote by $Y = ((s, y), (\sigma, \eta))$ the points in $T^*\widetilde{C}$. We also denote by Δ the diagonal in $T^*\widetilde{C} \times T^*\widetilde{C}$ and we will use the notation ${}_{\widetilde{C}}\Gamma$ and $\Gamma_{\widetilde{C}}$ introduced in (5-5).

Theorem 5.3. *Let* $\lambda^\pm : \mathscr{H}(\widetilde{C}) \to \mathscr{H}(\widetilde{C})$ *and*

$$\Lambda^\pm := (\rho \circ E_0)^* \circ \lambda^\pm \circ (\rho \circ E_0).$$

Then:

(1) $\Lambda^\pm \in \mathscr{D}'(M_0 \times M_0)$.

(2) *If*

(i) $\mathrm{WF}(\lambda^\pm)' \cap \{(Y_1, Y_2) : \pm\sigma_1 < 0$ *or* $\pm\sigma_2 < 0\} = \varnothing$,

(ii) $\mathrm{WF}(\lambda^+ - \lambda^-)' \cap \{(Y_1, Y_2) : \sigma_1$ *and* $\sigma_2 \neq 0\} \subset \Delta$,

then

(iii) $\mathrm{WF}(\lambda^\pm)' \cap \{(Y_1, Y_2) : \pm\sigma_1 > 0$ *and* $\pm\sigma_2 > 0\} \subset \Delta$.

(3) *Assume moreover that* $\lambda^\pm : \mathscr{H}(\widetilde{C}) \to \mathscr{H}(\widetilde{C})$ *and* ${}_{\widetilde{C}}\mathrm{WF}(\lambda^\pm)' = \mathrm{WF}(\lambda^\pm)'_{\widetilde{C}} = \varnothing$. *Then, if* (i) *and* (iii) *in* (2) *hold*, $\Lambda^\pm$ *satisfy* (μsc).

Proof. To prove (1) it suffices to check that $\rho \circ E_0 : \mathscr{D}(M_0) \to \mathscr{H}(\widetilde{C})$. If $\chi \in C_0^\infty(M_0)$ then, by Lemma 2.1, $\rho \circ E_0\chi = \rho \circ \chi_1 E\chi$ for some $\chi_1 \in C_0^\infty(M)$. Since $E : \mathscr{D}(M) \to \mathscr{E}(M)$ and $\rho : \mathscr{D}(M) \to \mathscr{H}(\widetilde{C})$ are continuous, this proves (1).

To prove (2) we write

$$\begin{aligned}\mathrm{WF}(\lambda^{\pm})' \cap \{\pm\sigma_1 > 0,\ \pm\sigma_2 > 0\} \\ &\subset (\mathrm{WF}(\lambda^{\mp})' \cap \{\pm\sigma_1 > 0,\ \pm\sigma_2 > 0\}) \cup (\mathrm{WF}(\lambda^{+} - \lambda^{-})' \cap \{\pm\sigma_1 > 0,\ \pm\sigma_2 > 0\}) \\ &\subset (\mathrm{WF}(\lambda^{\mp})' \cap \{\pm\sigma_1 > 0,\ \pm\sigma_2 > 0\}) \cup (\mathrm{WF}(\lambda^{+} - \lambda^{-})' \cap \{\sigma_1, \sigma_2 \neq 0\}).\end{aligned}$$

The first set in the last line is empty by (i), and the second is contained in Δ by (ii).

To prove (3) we follow an argument due to Moretti [2008]. We treat only the case of λ^+, the case of λ^- being similar, and omit the $+$ superscript. Let $\chi_i \in C_0^\infty(M_0)$, $i = 1, 2$. By Proposition 2.3 there exists $\psi_i \in C_0^\infty(C)$ (and hence $\psi_i \equiv 0$ near p) such that any null geodesic starting from $\operatorname{supp}\chi_i$ intersects C in $\{\psi_i = 1\}$. We have:

$$\begin{aligned}\chi_1\Lambda\chi_2 = {}& \chi_1(\rho\circ E)^*\psi_1\circ\lambda\circ\psi_2(\rho\circ E)\chi_2 + \chi_1(\rho\circ E)^*\psi_1\circ\lambda\circ(1-\psi_2)(\rho\circ E)\chi_2 \\ &+ \chi_1(\rho\circ E)^*(1-\psi_1)\circ\lambda\circ\psi_2(\rho\circ E)\chi_2 + \chi_1(\rho\circ E)^*(1-\psi_1)\circ\lambda\circ(1-\psi_2)(\rho\circ E)\chi_2 \\ =:{}& \Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4.\end{aligned}$$

By the properties of χ_i and ψ_i, we can find $\tilde{\chi}_i \in C_0^\infty(M)$ supported near p such that

(a) $(1-\psi_i)(\rho\circ E)\chi_i = (1-\psi_i)\rho\circ\tilde{\chi}_i E\chi_i$,

(b) no null geodesic from $\operatorname{supp}\chi_i$ intersects $\operatorname{supp}\tilde{\chi}_i$.

It follows from (b) and (5-6) that $\tilde{\chi}_i E\chi_i$ has a smooth, compactly supported kernel, hence

$$\tilde{\chi}_i E\chi_i : \mathscr{D}'(M) \to \mathscr{D}(M).$$

Since $(1-\psi_i)\rho : \mathscr{D}(M) \to \mathscr{H}(\widetilde{C})$, we see that

$$(1-\psi_i)\rho\circ E\chi_i : \mathscr{D}'(M) \to \mathscr{H}(\widetilde{C}), \tag{5-9}$$

hence

$$\chi_i(\rho\circ E)^*(1-\psi_i) : \mathscr{H}'(\widetilde{C}) \to \mathscr{D}(M). \tag{5-10}$$

It remains to examine the properties of $\psi_i(\rho\circ E)\chi_i$. By Proposition 5.2, $\psi_i(\rho\circ E)\chi_i : \mathscr{D}'(M) \to \mathscr{E}'(\widetilde{C})$. Since $\mathscr{E}'(\widetilde{C}) \subset \mathscr{H}'(\widetilde{C})$ continuously, we have

$$\psi_i(\rho\circ E)\chi_i : \mathscr{D}'(M) \to \mathscr{H}'(\widetilde{C}), \tag{5-11}$$

hence

$$\chi_i(\rho\circ E)^*\psi_i : \mathscr{H}(\widetilde{C}) \to \mathscr{D}(M). \tag{5-12}$$

From (5-9)–(5-12) and the assumption that $\lambda : \mathscr{H}(\widetilde{C}) \to \mathscr{H}(\widetilde{C})$ it follows that $\Lambda_i : \mathscr{D}'(M_0) \to \mathscr{D}(M_0)$, which hence has a smooth kernel for $i = 2, 3, 4$, and $\mathrm{WF}(\chi_1\Lambda\chi_2)' = \mathrm{WF}(\Lambda_1)'$.

To bound $\mathrm{WF}(\Lambda_1)'$ we choose $\tilde{\psi}_i \in C_0^\infty(\widetilde{C})$ such that $\tilde{\psi}_i\psi_i = \psi_i$ and write

$$\Lambda_1 = (\chi_1(\rho\circ E)\psi_1)\circ(\tilde{\psi}_1\lambda\tilde{\psi}_2)\circ(\psi_2(\rho\circ E)_2) =: K_1^*\circ d\circ K_2,$$

where $K_i = \psi_i(\rho \circ E)\chi_i \in \mathscr{E}'(M \times \widetilde{C})$ and $d = \tilde{\psi}_1 c \tilde{\psi}_2 \in \mathscr{E}'(\widetilde{C} \times \widetilde{C})$. The distributions K_1, K_2 and d have compact support. Moreover, we have

$$\mathrm{WF}(d)'_{\widetilde{C}} = {}_{\widetilde{C}}\mathrm{WF}(d)' = \mathrm{WF}(K_1)'_M = {}_M\mathrm{WF}(K_2^*)' = \varnothing.$$

In fact, the first two equalities follow from the corresponding hypothesis on $\mathrm{WF}(c)'$ and the last two from (5-8). We can then apply the results in [Hörmander 1990a, Chapter 8] on the composition of kernels and obtain that $K_2^* \circ d \circ K_1$ is well defined and

$$\mathrm{WF}(K_2^* \circ d \circ K_1) \subset \mathrm{WF}(K_2^*)' \circ \mathrm{WF}(d)' \circ \mathrm{WF}(K_1)'.$$

Now we apply Proposition 5.2(1), the fact that $\mathrm{WF}(d)' \subset \mathrm{WF}(\lambda)'$ and Lemma 5.1(1). We obtain that, if $(X_1, X_2) \in \mathrm{WF}(\Lambda)'$, necessarily $X_1, X_2 \in \mathcal{N}_+$ and $X_1 \sim X_2$, which is exactly condition (μsc). □

6. Pseudodifferential calculus

In this section we collect rather standard results on the pseudodifferential calculus on $\widetilde{C} = \mathbb{R} \times \mathbb{S}^{d-1}$. We will however need to consider *bihomogeneous* symbols on $\mathbb{R} \times \mathbb{S}^{d-1}$, i.e., symbols having different homogeneities in the covariables σ and η, dual to s and θ.

The reason for this is that the charge $q = -2D_s$ is not an elliptic differential operator in the usual sense (considered on $\widetilde{C}$), hence operators like $(q - z)^{-1}$ for $z \in \mathbb{C}\backslash\mathbb{R}$ are not in the usual pseudodifferential classes.

For $k, k' \in \mathbb{R}$, we denote by $H^k(\mathbb{R})$ and $H^{k'}(\mathbb{S}^{d-1})$ the Sobolev spaces on $\mathbb{R}$ and $\mathbb{S}^{d-1}$ of orders k and k' and by $\|\cdot\|_k$ and $\|\cdot\|_{k'}$ their respective norms. Furthermore, we denote by $H^{k,k'}(\mathbb{R} \times \mathbb{S}^{d-1})$ the Sobolev space on $\mathbb{R} \times \mathbb{S}^{d-1}$ of biorder (k, k'), that is, the completion of $C_0^\infty(\mathbb{R} \times \mathbb{S}^{d-1})$ for the norm

$$\|\psi\|_{k,k'} := \|\langle D_s\rangle^k \langle D_\theta\rangle^{k'} \psi\|_2.$$

We set also, for $p \in \mathbb{R}$,

$$B^p(\mathbb{R}) = \bigcap_{k \in \mathbb{R}} B(H^k(\mathbb{R}), H^{k-p}(\mathbb{R})),$$

equipped with its natural topology.

6A. *Pseudodifferential operators on* $\mathbb{R} \times \mathbb{R}^{d-1}$.

Definition 6.1. Let $p_1, p_2 \in \mathbb{R}$.

(1) We denote by $S^{p_1,p_2}(\mathbb{R} \times \mathbb{R}^{d-1})$ the space of symbols $a \in C^\infty(T^*\mathbb{R} \times T^*\mathbb{R}^{d-1})$ such that

$$|\partial_s^{\alpha_1} \partial_\sigma^{\beta_1} \partial_y^{\alpha_2} \partial_\eta^{\beta_2} a| \in O(\langle\sigma\rangle^{p_1 - |\beta_1|} \langle\eta\rangle^{p_2 - |\beta_2|}), \quad \alpha_1, \beta_1 \in \mathbb{N},\ \alpha_2, \beta_2 \in \mathbb{N}^{d-1}.$$

(2) We denote by $B^{p_1}S^{p_2}(\mathbb{R} \times \mathbb{S}^{d-1})$ the space of $a \in C^\infty(T^*\mathbb{R}^{d-1}, B^{p_1}(\mathbb{R}))$ such that

$$\|\partial_y^{\alpha_2} \partial_\eta^{\beta_2} a\|_{p_1,k_1} \in O(\langle\eta\rangle^{p_2 - |\beta_2|}), \quad \alpha_2, \beta_2 \in \mathbb{N}^{d-1},$$

where $\|\cdot\|_{p_1,k_1}$ is any seminorm of a in $B^{p_1}(\mathbb{R})$.

Using the Weyl quantization on $\mathbb{R}\times\mathbb{R}^{d-1}$, we obtain a map

$$S^{p_1,p_2}(\mathbb{R}\times\mathbb{R}^{d-1}) \to B(C_0^\infty(\mathbb{R}\times\mathbb{R}^{d-1}), C^\infty(\mathbb{R}\times\mathbb{R}^{d-1})), \quad a\mapsto \mathrm{Op}(a),$$

whose range, denoted by $\Psi^{p_1,p_2}(\mathbb{R}\times\mathbb{R}^{d-1})$, is the space of pseudodifferential operators on $\mathbb{R}\times\mathbb{R}^{d-1}$ of biorder (p_1, p_2). Similarly, using the Weyl quantization on $\mathbb{R}^{d-1}$, we obtain a map

$$B^{p_1}S^{p_2}(\mathbb{R}\times\mathbb{R}^{d-1}) B(C_0^\infty(\mathbb{R}\times\mathbb{R}^{d-1}), C^\infty(\mathbb{R}\times\mathbb{R}^{d-1})), \quad a\mapsto \mathrm{Op}(a),$$

whose range will be denoted by $B^{p_1}\Psi^{p_2}(\mathbb{R}\times\mathbb{R}^{d-1})$.

6B. ***Pseudodifferential operators on $\widetilde{C}$.*** Let $A : C_0^\infty(\widetilde{C}) \to C^\infty(\widetilde{C})$. If $\chi_i \in C^\infty(\mathbb{S}^{d-1})$, $i = 1, 2$, are cutoff functions supported in chart open sets $\Omega_i \subset \mathbb{S}^{d-1}$ and $\phi_i : \Omega_i \to \mathbb{R}^{d-1}$ are coordinate charts, then $\phi_1^*\circ\chi_1 A\chi_2\circ\phi_2^{-1*} : C_0^\infty(\mathbb{R}\times\mathbb{R}^{d-1}) \to C^\infty(\mathbb{R}\times\mathbb{R}^{d-1})$.

Definition 6.2. (1) We denote by $\Psi^{p_1,p_2}(\widetilde{C})$ the space of operators $A : C_0^\infty(\widetilde{C}) \to C^\infty(\widetilde{C})$ such that, for any χ_i and ϕ_i as above, $\phi_1^*\circ\chi_1 A\chi_2\circ\phi_2^{-1*} \in \Psi^{p_1,p_2}(\mathbb{R}\times\mathbb{R}^{d-1})$.

(2) We denote by $B^{p_1}\Psi^{p_2}(\widetilde{C})$ the space of operators $A : C_0^\infty(\widetilde{C}) \to C^\infty(\widetilde{C})$ such that, for any χ_i and ϕ_i as above, $\phi_1^*\circ\chi_1 A\chi_2\circ\phi_2^{-1*} \in B^{p_1}\Psi^{p_2}(\mathbb{R}\times\mathbb{R}^{d-1})$.

(3) We set

$$\Psi^{-\infty,p_2}(\widetilde{C}) = \bigcap_{p_1\in\mathbb{R}} \Psi^{p_1,p_2}(\widetilde{C}) \quad\text{and}\quad B^{-\infty}\Psi^{p_2}(\widetilde{C}) = \bigcap_{p_1\in\mathbb{R}} B^{p_1}\Psi^{p_2}(\widetilde{C}).$$

(4) We set

$$\widetilde{\Psi}^{p_1,p_2}(\widetilde{C}) = \Psi^{p_1,p_2}(\widetilde{C}) + B^{-\infty}\Psi^{p_2}(\widetilde{C}).$$

Note that if one defines, analogously, $\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) := \bigcap_{p_1\in\mathbb{R}} \widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$, then actually $\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) = B^{-\infty}\Psi^{p_2}(\widetilde{C})$. Moreover, it is easy to check that

$$\widetilde{\Psi}^{p_1,p_2}(\widetilde{C}) \circ \widetilde{\Psi}^{q_1,q_2}(\widetilde{C}) \subset \widetilde{\Psi}^{p_1+p_2,q_1+q_2}(\widetilde{C}).$$

We refer the reader to [Rodino 1975; Borsero and Schulz 2014; Ruzhansky and Turunen 2010] and references therein for more details on the pseudodifferential calculus on products of manifolds.[4]

6C. ***The Beals criterion.*** Let us denote by $\Psi^p(\mathbb{S}^{d-1})$ the classes of standard pseudodifferential operators on $\mathbb{S}^{d-1}$. It is well known that $\Psi^p(\mathbb{S}^{d-1})$ can be characterized by the Beals criterion, namely an operator $A : C^\infty(\mathbb{S}^{d-1}) \to C^\infty(\mathbb{S}^{d-1})$ belongs to $\Psi^p(\mathbb{S}^{d-1})$ if and only if

$$\mathrm{ad}_{f_1}\cdots\mathrm{ad}_{f_n}\,\mathrm{ad}_{X_1}\cdots\,\mathrm{ad}_{X_m} A : H^k(\mathbb{S}^{d-1}) \to H^{k-p+n}(\mathbb{S}^{d-1}), \quad n, m\in\mathbb{N},\ k\in\mathbb{Z}, \tag{6-1}$$

for any $f_i \in C^\infty(\mathbb{S}^{d-1})$ and smooth vector fields X_j on $\mathbb{S}^{d-1}$ [Ruzhansky and Turunen 2010]. Moreover, one can find a finite set of such f_i and X_j such that the topology on $\Psi^p(\mathbb{S}^{d-1})$ given by the collection of the norms of the multicommutators is equivalent to the standard topology on $\Psi^p(\mathbb{S}^{d-1})$, given by the

[4] Note however that the literature discusses mostly the case when both manifolds are compact.

symbol space topologies of the pullbacks $\phi_i^* \circ \chi_i A \chi_j \circ \phi_j$ in Definition 6.2, for a fixed covering of $\mathbb{S}^{d-1}$ by chart neighborhoods U_i.

These characterizations immediately carry over to the classes $B^{p_1}\Psi^{p_2}(\widetilde{C})$. In fact it is easy to see that $A \in B^{p_1}S^{p_2}(\widetilde{C})$ if and only if

$$\mathrm{ad}_{f_1}\cdots\mathrm{ad}_{f_n}\,\mathrm{ad}_{X_1}\cdots\mathrm{ad}_{X_m}A : H^{k,k'}(\widetilde{C}) \to H^{k-p_1,k'-p_2+n}(\mathbb{S}^{d-1}), \quad n, m \in \mathbb{N},\ k, k' \in \mathbb{Z}. \tag{6-2}$$

This result can be deduced from the previous one by considering the operators

$$((u_1| \otimes \mathbb{1}_{\mathbb{S}^{d-1}}) \circ A \circ (|u_2) \otimes \mathbb{1}_{\mathbb{S}^{d-1}}) : C^\infty(\mathbb{S}^{d-1}) \to C^\infty(\mathbb{S}^{d-1})$$

for $u_1 \in H^{-k+p_1}(\mathbb{R})$ and $u_2 \in H^k(\mathbb{R})$, which belong to $\Psi^{p_2}(\mathbb{S}^{d-1})$ if (6-2) holds. Applying the result recalled above about the equivalence of the standard topology and the topology given by the multicommutator norms, one obtains that $A \in B^{p_1}\Psi^{p_2}(\widetilde{C})$ if (6-2) holds.

In the usual case one can deduce from the Beals criterion standard results on the functional calculus for pseudodifferential operators, for example on complex powers of elliptic pseudodifferential operators [Bony 1997]. These results are easy to extend to the classes $B^{p_1}\Psi^{p_2}(\widetilde{C})$. We will need only a very simple one, which we now state. Recall that $\widetilde{\Psi}^{-\infty,0}(\widetilde{C}) = B^{-\infty}\Psi^0(\widetilde{C}) \subset B(L^2(\widetilde{C}))$. The spectrum of $b \in B(L^2(\widetilde{C}))$ is denoted by $\mathrm{spec}(b)$.

Proposition 6.3. *Let $b \in \widetilde{\Psi}^{-\infty,0}(\widetilde{C})$ and let be F holomorphic near* $\mathrm{spec}(b)$ *with $F(0) = 0$. Then $F(b) \in \widetilde{\Psi}^{-\infty,0}(\widetilde{C})$.*

Proof. The proof consists of expressing $F(b)$ as a contour integral and applying the Beals criterion to the resolvent $(b - z)^{-1}$. □

6D. *Essential support.* We denote by $\Psi^p_{\mathrm{ph}}(\mathbb{R})$, $p \in \mathbb{R}$, the class of global pseudodifferential operators on $\mathbb{R}$ with polyhomogeneous symbols.

Definition 6.4. The *essential support* of $a \in \Psi^{p_1,p_2}(\widetilde{C})$, denoted by $\mathrm{ess\,supp}(a) \subset T^*\mathbb{R}\backslash Z$, is defined by $(s_0, \sigma_0) \notin \mathrm{ess\,supp}(a)$ if there exists $b \in \Psi^0_{\mathrm{ph}}(\mathbb{R})$ that is elliptic at (s_0, σ_0) such that $b \circ a \in \Psi^{-\infty,p_2}(\widetilde{C})$.

Clearly $\mathrm{ess\,supp}(a)$ is a closed conic subset of $T^*\mathbb{R}\backslash Z$. Moreover, one can equivalently require that $a \circ b \in \Psi^{-\infty,p_2}(\widetilde{C})$ for some $b \in \Psi^0_{\mathrm{ph}}(\mathbb{R})$ that is elliptic at (s_0, σ_0).

6E. *Wavefront set of kernels.* For $N = \mathbb{R}, \mathbb{S}^{d-1}, \mathbb{R}\times\mathbb{S}^{d-1}$, we denote by Δ_N the diagonal in $T^*N \times T^*N$ and by Z_N the zero section in T^*N.

For an operator $a \in \Psi^{p_1,p_2}(\mathbb{R} \times \mathbb{S}^{d-1})$, it is in general not true that $\mathrm{WF}(a)'$ is contained in the full diagonal $\Delta_{\mathbb{R}\times\mathbb{S}^{d-1}}$ (as would be the case for an operator in $\Psi^p(\mathbb{R} \times \mathbb{S}^{d-1})$). Instead one has the following estimate, which can be thought as a natural generalization of the usual estimate for the wavefront set of tensor products of distributions (in this case Schwartz kernels) [Borsero and Schulz 2014].

Lemma 6.5. *Let $a \in \Psi^{p_1,p_2}(\mathbb{R} \times \mathbb{S}^{d-1})$. Then*

$$\mathrm{WF}(a)' \subset \Delta_{\mathbb{R}} \times \Delta_{\mathbb{S}^{d-1}} \cup \Delta_{\mathbb{R}} \times (Z_{\mathbb{S}^{d-1}} \times Z_{\mathbb{S}^{d-1}}) \cup (Z_{\mathbb{R}} \times Z_{\mathbb{R}}) \times \Delta_{\mathbb{S}^{d-1}}.$$

Less precise estimates are valid for the $\widetilde{\Psi}^{p_1,p_2}(\mathbb{R} \times \mathbb{S}^{d-1})$ classes:

Lemma 6.6. (1) *Let* $a \in B^{-\infty}\Psi^{p_2}(\widetilde{C})$. *Then*

$$\mathrm{WF}(a)' \cap \{(Y_1, Y_2) : \sigma_1 \neq 0 \ \text{or} \ \sigma_2 \neq 0\} = \varnothing.$$

(2) *Let* $a \in \widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$. *Then*

$$_{\widetilde{C}}\mathrm{WF}(a)' = \mathrm{WF}(a)'_{\widetilde{C}} = \varnothing.$$

The proof is given in Appendix A5.

6F. ***Toeplitz pseudodifferential operators on*** $\widetilde{\boldsymbol{C}}$**.** We recall that $\mathscr{H}(\widetilde{C}) = \bigcap_{m\in\mathbb{R}} H^m(\widetilde{C}) = \bigcap_{k\in\mathbb{R}} H^{k,k}(\widetilde{C})$. Let us set

$$L^2_{\pm}(\widetilde{C}) := \mathbb{1}_{\mathbb{R}^{\pm}}(D_s)L^2(\widetilde{C})$$

and denote by $i_{\pm} : L^2_{\pm}(\widetilde{C}) \to L^2(\widetilde{C})$ the corresponding isometric injection, so that $\pi_{\pm} := i_{\pm}i^*_{\pm} = \mathbb{1}_{\mathbb{R}^{\pm}}(D_s)$ is the orthogonal projection on $L^2_{\pm}(\widetilde{C})$ in $L^2(\widetilde{C})$. We also set

$$\mathscr{H}_{\pm}(\widetilde{C}) := i^*_{\pm}\mathscr{H}(\widetilde{C}) \subset \mathscr{H}(\widetilde{C}). \tag{6-3}$$

We will see in Section 7 that this provides a useful setup for the discussion of the positivity condition $\lambda^{\pm} \geq 0$ for the two-point functions of a Hadamard state.

Writing $\mathbb{1}_{\mathbb{R}^{\pm}} = \chi\mathbb{1}_{\mathbb{R}^{\pm}} + (1-\chi)\mathbb{1}_{\mathbb{R}^{\pm}}$ for a cutoff function $\chi \in C_0^{\infty}(\mathbb{R})$ equal to 1 near 0, we see that

$$\pi_{\pm} \in \widetilde{\Psi}^{0,0}(\widetilde{C}). \tag{6-4}$$

For $\alpha, \beta \in \{+, -\}$ and $p_1, p_2 \in \mathbb{R}$, we set

$$\widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C}) := i_{\alpha} \circ \widetilde{\Psi}^{p_1,p_2}(\widetilde{C}) \circ i^*_{\beta}.$$

By (6-3) we see that $\widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C}) : \mathscr{H}_{\beta}(\widetilde{C}) \to \mathscr{H}_{\alpha}(\widetilde{C})$. Moreover, if we set

$$R_{\alpha\beta} : \widetilde{\Psi}^{p_1,p_2}(\widetilde{C}) \to \widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C}), \quad a \mapsto i^*_{\alpha} \circ a \circ i_{\beta}$$

then, using (6-4), we see that $R_{\alpha\beta}$ has right inverse

$$T_{\alpha\beta} : \widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C}) \to \widetilde{\Psi}^{p_1,p_2}(\widetilde{C}), \quad a \mapsto i_{\alpha} \circ a \circ i^*_{\beta},$$

which allows us to identify $\widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C})$ with $\operatorname{Ran} T_{\alpha\beta} \subset \widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$. From (6-4) we also have

$$\widetilde{\Psi}^{p_1,p_2}_{\alpha\beta}(\widetilde{C}) \circ \widetilde{\Psi}^{q_1,q_2}_{\beta\gamma}(\widetilde{C}) \subset \widetilde{\Psi}^{p_1+q_1,p_2+q_2}_{\alpha\gamma}(\widetilde{C}). \tag{6-5}$$

7. Construction of Hadamard states on the cone

From the discussion in Section 5B, in particular Theorem 5.3, we are led to the following definition:

Definition 7.1. A pair of maps $\lambda^{\pm} : \mathscr{H}(\widetilde{C}) \to \mathscr{H}(\widetilde{C})$ is called a pair of *Hadamard two-point functions* on the cone C if

$$ {}_{\widetilde{C}}\mathrm{WF}(\lambda^{\pm})' = \mathrm{WF}(\lambda^{\pm})'_{\widetilde{C}} = \varnothing, \tag{Had-i}$$

$$ \mathrm{WF}(\lambda^{\pm})' \cap \{(Y_1, Y_2) : \pm\sigma_1 < 0 \ \text{or} \ \pm\sigma_2 < 0\} = \varnothing, \tag{Had-ii}$$

$$ \lambda^{+} - \lambda^{-} = 2D_s, \tag{Had-iii}$$

$$ \lambda^{\pm} \geq 0 \quad \text{on } \mathscr{H}(\widetilde{C}). \tag{Had-iv}$$

As the name suggests, if $\lambda^{\pm}$ are Hadamard two-point functions on C in the sense of the above definition, then $\Lambda^{\pm}$ defined in (4-4) are Hadamard two-point functions on M_0 (as follows from Theorem 5.3).

We now discuss in more detail the various conditions in (Had-i)–(Had-iv). It is natural to consider pseudodifferential two-point functions, i.e., to assume that $\lambda^{\pm} \in \widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$. Moreover to analyze conditions (Had-iii)–(Had-iv) it is convenient to reduce oneself to $\lambda^{\pm}$ of the form

$$ \lambda^{\pm} = (2|D_s|)^{1/2} c^{\pm} (2|D_s|)^{1/2}, \quad \text{where} \quad c^{\pm} \in \widetilde{\Psi}^{p_1,p_2}(\widetilde{C}), \tag{7-1}$$

for $p_1, p_2 \in \mathbb{R}$. Note that, writing $(2|D_s|)^{1/2}$ as $\chi(D_s)(2|D_s|)^{1/2} + (1-\chi(D_s))(2|D_s|)^{1/2}$ for $\chi \in C_0^{\infty}(\mathbb{R})$ equal to 1 near 0, we see that (7-1) implies that $\lambda^{\pm} \in \widetilde{\Psi}^{p_1+1,p_2}(\widetilde{C})$.

7A. *Wavefront set.* We first analyze conditions (Had-i)–(Had-ii).

Proposition 7.2. *Assume that*

$$ \lambda^{\pm} = a^{\pm} + r^{\pm}, \quad a^{\pm} \in \Psi^{p_1,p_2}(\widetilde{C}), \quad r^{\pm} \in \widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}), \quad (\mathbb{R} \times \mathbb{R}^{\mp}) \cap \operatorname{ess\,supp}(a^{\pm}) = \varnothing. \tag{7-2}$$

Then $\lambda^{\pm}$ satisfies conditions (Had-i)–(Had-ii).

Proof. The fact that $\lambda^{\pm}$ satisfy (Had-i) follows from Lemma 6.6(2). Also, since, by Lemma 6.6(1), $r^{\pm}$ satisfy (Had-ii) we can assume that $\lambda^{\pm} = a^{\pm}$. We treat only the case of λ^{+} and use the notation in the proof of Lemma 6.6. Let $\widetilde{Y}_1, \widetilde{Y}_2 \in T^*\widetilde{C} \backslash Z$ with $\tilde{\sigma}_1 \neq 0$ or $\tilde{\sigma}_2 \neq 0$. Let us assume that $\tilde{\sigma}_1 \neq 0$, the case $\tilde{\sigma}_2 \neq 0$ being similar, using the remark after Definition 6.4.

Since $(\mathbb{R} \times \mathbb{R}^{+}) \cap \operatorname{ess\,supp}(a^{+}) = \varnothing$, we can find a cutoff function χ_1 with $\chi_1(\tilde{s}_1) \neq 0$, a neighborhood V_1 of $\tilde{\sigma}_1$ and some $m_1 \in \Psi^{0}_{\mathrm{ph}}(\mathbb{R})$ elliptic at $(\tilde{s}_1, \tilde{\sigma}_1)$ such that $(1-m_1)(s, D_s)v_{\sigma,\lambda} \in O(\langle\lambda\rangle^{-\infty})$ in all $H^k(\mathbb{R})$ and $m_1(s, D_s) \circ a \in \widetilde{\Psi}^{-\infty,p_2}(\widetilde{C})$. The fact that $(\widetilde{Y}_1, \widetilde{Y}_2) \notin \mathrm{WF}(a)'$ then follows from the same arguments as in the proof of Lemma 6.6. □

In terms of $c^{\pm}$ appearing in (7-1), a natural condition implying (7-2) is

$$ \mathbb{1}_{\mathbb{R}^{\mp}}(D_s) c^{\pm} \in \widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}), \tag{$\mu\mathrm{sc}_C$}$$

which clearly implies that $\lambda^{\pm}$ satisfy (7-2).

Lemma 7.3. *Let $\lambda^{\pm}$ be given by* (7-1) *and such that* ($\mu\mathrm{sc}_C$) *holds. Then*

$$ c^{\pm} = \mathbb{1}_{\mathbb{R}^{\pm}}(D_s) + \widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}).$$

Proof. In terms of $c^{\pm}$, (Had-iii) becomes $c^{+}-c^{-}=\operatorname{sgn}(D_s)$. Let $\chi^{\pm}\in C^{\infty}(\mathbb{R})$ be cutoff functions equal to 1 near $\pm\infty$ and to 0 near $\mp\infty$. From ($\mu\mathrm{sc}_C$) and pseudodifferential calculus we obtain that

$$c^{\pm}=\chi^{\pm}(D_s)c^{\pm}\chi^{\mp}(D_s)+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}). \tag{7-3}$$

Using successively (7-3) and $c^{+}-c^{-}=\operatorname{sgn}(D_s)$, we obtain

$$\begin{aligned} c^{\pm} &= \chi^{\pm}c^{\pm}\chi^{\pm}+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) \\ &= \chi^{\pm}(c^{\mp}\pm\operatorname{sgn}(D_s))\chi^{\pm}+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) \\ &= \chi^{\pm}c^{\mp}\chi^{\pm}+\chi^{\pm}\chi^{\pm}+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) \\ &= \chi^{\pm}\chi^{\mp}c^{\mp}\chi^{\mp}\chi^{\pm}+\chi^{\pm}+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) \\ &= \chi^{\pm}+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}) \\ &= \mathbb{1}_{\mathbb{R}^{\pm}}(D_s)+\widetilde{\Psi}^{-\infty,p_2}(\widetilde{C}). \end{aligned}$$

□

7B. *Positivity.* We now discuss conditions (Had-iii)–(Had-iv). In terms of $c^{\pm}$ they become

$$c^{+}-c^{-}=\operatorname{sgn}(D_s), \tag{7-4-iii}$$
$$c^{\pm}\geq 0 \quad \text{on } \mathscr{H}(\widetilde{C}). \tag{7-4-iv}$$

To analyze (7-4-iii)–(7-4-iv) we use the framework of Section 6F. We denote c^{+} simply by c and set

$$c_{\alpha\beta}=i_{\alpha}^{*}\circ c\circ i_{\beta},\quad \alpha,\beta\in\{+,-\},$$

so that

$$c=\sum_{\alpha,\beta\in\{+,-\}} i_{\alpha}c_{\alpha\beta}i_{\beta}^{*}. \tag{7-5}$$

Then (7-4-iii)–(7-4-iv) is equivalent to

$$\begin{pmatrix} c_{++} & c_{+-} \\ c_{-+} & c_{--}\end{pmatrix}\geq 0 \quad\text{and}\quad \begin{pmatrix} c_{++}-\mathbb{1} & c_{+-} \\ c_{-+} & c_{--}+\mathbb{1}\end{pmatrix}\geq 0 \quad \text{on } \mathscr{H}_{+}(\widetilde{C})\oplus\mathscr{H}_{-}(\widetilde{C}), \tag{7-6}$$

which is equivalent to

(i) $c_{++}\geq 0$, $c_{--}\geq\mathbb{1}$ and $c_{-+}=c_{+-}^{*}$.

(ii) $$\begin{aligned} \left|(u_{+}\mid c_{+-}u_{-})\right| &\leq (u_{+}\mid c_{++}u_{+})^{1/2}(u_{-}\mid c_{--}u_{-})^{1/2}, \\ \left|(u_{+}\mid c_{+-}u_{-})\right| &\leq (u_{+}\mid (c_{++}-\mathbb{1})u_{+})^{1/2}(u_{-}\mid (c_{--}+\mathbb{1})u_{-})^{1/2}, \end{aligned}\quad u_{\pm}\in\mathscr{H}_{\pm}(\widetilde{C}).$$

Condition (ii) above is implied by

$$\left|(u_{+}\mid c_{+-}u_{-})\right|\leq (u_{+}\mid (c_{++}-\mathbb{1})u_{+})^{1/2}(u_{-}\mid c_{--}u_{-})^{1/2},\quad u_{\pm}\in\mathscr{H}_{\pm}(\widetilde{C}).$$

We are now in position to prove the following theorem, which is the analog of [Gérard and Wrochna 2014, Theorem 7.5] in the present situation. It provides a rather large class of Hadamard two-point functions on C and hence, by Theorem 5.3, of Hadamard states on M_0.

Theorem 7.4. *Assume that*

$$c_{++} = \mathbb{1} + a_+^* a_+, \quad c_{--} = a_-^* a_-, \quad c_{+-} = c_{-+}^* = a_+^* d a_-$$

for $a_+ \in \widetilde{\Psi}_{++}^{-\infty,0}(\widetilde{C})$, $a_- \in \widetilde{\Psi}_{--}^{-\infty,0}(\widetilde{C})$, $d \in \widetilde{\Psi}_{+-}^{0,0}(\widetilde{C})$ *with* $\|d\|_{B(L^2_-(\widetilde{C}),L^2_+(\widetilde{C}))} \le \mathbb{1}$.

Let c be given by (7-5), $\lambda^+ = (2|D_s|)^{1/2} c (2|D_s|)^{1/2}$ *and* $\lambda^- = \lambda^+ - 2D_s$. *Then* $\lambda^\pm$ *is a pair of Hadamard two-point functions on the cone.*

Proof. We set, as before, $\lambda^\pm = (2|D_s|)^{1/2} c^\pm (2|D_s|)^{1/2} \in \widetilde{\Psi}^{1,0}(\widetilde{C})$, so that $c^+ = c$ and $c^- = c - \mathrm{sgn}(D_s)$. Conditions (7-4-iii)–(7-4-iv) follow from the above discussion. It remains to check condition ($\mu\mathrm{sc}_C$). We embed the spaces $\widetilde{\Psi}_{\alpha\beta}^{p_1,p_2}(\widetilde{C})$ into $\widetilde{\Psi}^{p_1,p_2}(\widetilde{C})$ as explained at the end of Section 6F and we have

$$\begin{aligned} c^+ &= a_+^* a_+ + a_+^* d a_- + a_-^* d^* a_+ + a_-^* a_- + \mathbb{1}_{\mathbb{R}^+}(D_s), \\ c^- &= a_+^* a_+ + a_+^* d a_- + a_-^* d^* a_+ + a_-^* a_- + \mathbb{1}_{\mathbb{R}^-}(D_s), \end{aligned}$$

hence

$$\begin{aligned} \mathbb{1}_{\mathbb{R}^-}(D_s) c^+ &= a_+^* a_+ + a_+^* d a_- \in \widetilde{\Psi}^{-\infty,0}(\widetilde{C}), \\ \mathbb{1}_{\mathbb{R}^+}(D_s) c^- &= a_-^* d^* a_+ + a_-^* a_- \in \widetilde{\Psi}^{-\infty,0}(\widetilde{C}), \end{aligned}$$

and condition ($\mu\mathrm{sc}_C$) is satisfied. □

Remark 7.5. The special choice of vanishing a_+, a_- and d in Theorem 7.4 gives two-point functions

$$\lambda^\pm = \pm 2 \mathbb{1}_{\mathbb{R}^\pm}(D_s) D_s.$$

In the setting of asymptotically flat spacetimes with past time infinity i^-, these correspond to the Hadamard state found and further studied in [Moretti 2006; 2008].

8. Pure Hadamard states

In this section we first characterize *pure* Hadamard states on the cone C. We then prove that any pure Hadamard state ω_C on C induces a pure Hadamard state ω_0 in M_0.

8A. ***An abstract criterion for purity.*** Let $(\mathcal{Y}, \sigma)$ a complex symplectic space and ω a gauge invariant quasifree state on $\mathrm{CCR}(\mathcal{Y}, \sigma)$, with complex covariances $\lambda^\pm$.

Let $\mathcal{Y}^{\mathrm{cpl}}$ the completion of $\mathcal{Y}$ for the norm

$$\|y\|_\omega := (\bar{y} \cdot \lambda^+ y + \bar{y} \cdot \lambda^- y)^{1/2}. \tag{8-1}$$

Let us introduce the hermitian form $q = \mathrm{i}\sigma \in L_{\mathrm{h}}(\mathcal{Y}, \mathcal{Y}^*)$. Clearly q and $\lambda^\pm$ extend uniquely to $\mathcal{Y}^{\mathrm{cpl}}$. Then, by [Araki and Shiraishi 1971/72], ω is pure if and only if

(1) q is nondegenerate on $\mathcal{Y}^{\mathrm{cpl}}$,

(2) there exists an involution $\kappa : \mathcal{Y}^{\mathrm{cpl}} \to \mathcal{Y}^{\mathrm{cpl}}$ such that $\kappa^* q \kappa = q$, $q\kappa \ge 0$ and $\lambda^\pm = \frac{1}{2} q(\kappa \pm \mathbb{1})$.

From this discussion we immediately obtain the following lemma:

Lemma 8.1. *Let $(\mathcal{Y}_i, \sigma_i)$, $i = 1, 2$, be two complex symplectic spaces and $\rho : \mathcal{Y}_1 \to \mathcal{Y}_2$ an injective map such that $\rho^*\sigma_2\rho = \sigma_1$. Let ω_2 be a pure, gauge-invariant quasifree state on $\mathrm{CCR}(\mathcal{Y}_2, \sigma_2)$. Let ω_1 be the gauge-invariant, quasifree state on $\mathrm{CCR}(\mathcal{Y}_1, \sigma_1)$ defined by the complex covariances*

$$\lambda_1^{\pm} = \rho^* \lambda_2^{\pm} \rho.$$

Then, if $\rho\mathcal{Y}_1$ is dense in $\mathcal{Y}_2$ for the norm $\|\cdot\|_{\omega_2}$ defined in (8-1), *the state ω_1 is pure on* $\mathrm{CCR}(\mathcal{Y}_1, \sigma_1)$.

8B. ***Pure Hadamard states on the cone.*** The following theorem is the exact analog of [Gérard and Wrochna 2014, Theorem 7.10]. In what follows we will use the notations introduced in Section 6F.

Theorem 8.2. *Let $\lambda^{\pm}$ be the two-point functions of a state ω_C on $(\mathcal{H}(\widetilde{C}), \sigma_C)$ of the form* (7-1) *and satisfying $(\mu\mathrm{sc}_C)$. Then ω_C is pure if and only if there exists $a \in \widetilde{\Psi}^{-\infty,0}_{-+}(\widetilde{C})$ such that*

$$c^+ = \begin{pmatrix} \mathbb{1} + a^*a & a^*(\mathbb{1} + aa^*)^{1/2} \\ (\mathbb{1} + aa^*)^{1/2}a & aa^* \end{pmatrix}.$$

Proof. We consider the pair $c^{\pm}$ obtained from $\lambda^{\pm}$, write as before c^+ for c and identify c with the matrix

$$\begin{pmatrix} c_{++} & c_{+-} \\ c_{-+} & c_{--} \end{pmatrix}.$$

Arguing as in the proof of [Gérard and Wrochna 2014, Theorem 7.10], we obtain that the state ω_C on $(\mathcal{H}(\widetilde{C}), \sigma_C)$ with covariances $\lambda^{\pm}$ is pure if and only if

$$c = \begin{pmatrix} \mathbb{1} + a^*a & a^*(\mathbb{1} + aa^*)^{1/2} \\ (\mathbb{1} + aa^*)^{1/2}a & aa^* \end{pmatrix} \tag{8-2}$$

for some $a : L^2_+(\widetilde{C}) \to L^2_-(\widetilde{C})$. This proves the "if".

Let us now prove the "only if". Since we assumed that $c^{\pm} \in \widetilde{\Psi}^{0,0}(\widetilde{C})$ satisfy $(\mu\mathrm{sc}_C)$, we obtain that

$$a^*a \in \widetilde{\Psi}^{-\infty,0}_{++}(\widetilde{C}), \quad (\mathbb{1} + aa^*)^{1/2}a \in \widetilde{\Psi}^{-\infty,0}_{-+}(\widetilde{C}). \tag{8-3}$$

We claim that

$$(\mathbb{1} + aa^*)^{-1/2} \in \mathbb{1} + \widetilde{\Psi}^{-\infty,0}_{--}(\widetilde{C}). \tag{8-4}$$

Let us prove (8-4). We use the operators $R_{\alpha\beta}$ and $T_{\alpha\beta}$ defined at the end of Section 6F. We first embed aa^* into $\widetilde{\Psi}^{-\infty,0}(\widetilde{C})$, i.e., consider $b = T_{--}(a^*a)$. Then $b \geq 0$ on $L^2(\widetilde{C})$ and, applying Proposition 6.3 to $F(z) = (1+z)^{1/2} - 1$, we obtain that $(\mathbb{1} + b)^{-1/2} - \mathbb{1} \in \widetilde{\Psi}^{-\infty,0}(\widetilde{C})$. Writing b as a 2×2 matrix acting on $L^2_+(\widetilde{C}) \oplus L^2_-(\widetilde{C})$ we see that $R_{++}((\mathbb{1} + b)^{1/2}) = (\mathbb{1} + aa^*)^{1/2}$, which proves (8-4). From (8-4) and (8-3) we obtain that $a \in \widetilde{\Psi}^{-\infty,0}_{-+}(\widetilde{C})$. □

In the next lemma we identify the completion of $\mathcal{H}(\widetilde{C})$ for the norm (8-1) associated to any Hadamard state considered in Theorem 8.2.

Let us first fix some notation. For $a : L^2_+(\widetilde{C}) \to L^2_-(\widetilde{C})$ we denote by $c^+(a)$ the operator defined in (8-2) and set $c^-(a) = c^+(a) - \mathrm{sgn}(D_s)$ and

$$\lambda^{\pm}(a) = (2|D_s|)^{1/2} c^{\pm}(a) (2|D_s|)^{1/2}. \tag{8-5}$$

If $\mathscr{H}$ is a Hilbert space and $h \geq 0$ is a selfadjoint operator on $\mathscr{H}$ with $\operatorname{Ker} h = \{0\}$, we denote by $h\mathscr{H}$ the completion of $\operatorname{Dom} h^{-1}$ (the range of h) for the norm $\|h^{-1}u\|_{\mathscr{H}}$.

Lemma 8.3. *Let $a : L^2_+(\widetilde{C}) \to L^2_-(\widetilde{C})$. Then the completion of $\mathscr{H}(\widetilde{C})$ for the norm $(\,\cdot\,|\,(\lambda^+(a)+\lambda^-(a))\,\cdot\,)^{1/2}$ equals $|D_s|^{-1/2}L^2(\widetilde{C})$.*

Proof. By (8-5) and the definition of $|D_s|^{-1/2}L^2(\widetilde{C})$, it suffices to prove that the completion of $\mathscr{H}(\widetilde{C})$ for the norm $\bigl(u \bigm| (c^+(a)+c^-(a))u\bigr)^{1/2}$ equals $L^2(\widetilde{C})$. Let

$$u(a) = \begin{pmatrix} (\mathbb{1}+aa^*)^{1/2} & a \\ a^* & (\mathbb{1}+a^*a)^{1/2} \end{pmatrix}$$

and note that

$$u(a)^* c^{\pm}(0) u(a) = c^{\pm}(a). \tag{8-6}$$

Moreover, using the identity $af(a^*a) = f(aa^*)a$, valid for any Borel function f, we obtain that $u(a)^{-1} = u(-a)$, hence $u(a) : L^2(\widetilde{C}) \to L^2(\widetilde{C})$ is boundedly invertible. By (8-6), it suffices to treat the case $a = 0$, which is obvious since $c^+(0)+c^-(0) = \mathbb{1}$. □

8C. *Pure Hadamard states in M_0.* Our main result concerns the purity of the states induced in the bulk. We postpone the introduction of the key technical ingredients of the proof to Section 8D for the sake of self-consistency of our results on the characteristic Cauchy problem.

Theorem 8.4. *Assume that* $\dim M \geq 4$. *Let ω_C be a pure Hadamard state on* $\mathrm{CCR}(\mathscr{H}(\widetilde{C}), \sigma_C)$ *as in Theorem 8.2. Then the state ω induced by ω_C on* $\mathrm{CCR}(C_0^\infty(M_0)\,/\,PC_0^\infty(M_0), E_0)$ *is a pure state.*

Proof. The proof relies on Lemma 8.1 and on some results on the characteristic Cauchy problem in M_0, proved below in Section 8D. Recall that the map $\rho : \mathrm{Sol}_{\mathrm{sc}}(P_0) \to \mathscr{H}(\widetilde{C})$ was introduced in Definition 4.1. By Lemmas 8.1 and 8.3 it suffices to check that $\rho(\mathrm{Sol}_{\mathrm{sc}}(P_0))$ is dense in $|D_s|^{-1/2}L^2(\widetilde{C})$. Since $C_0^\infty(\mathbb{R}\times\mathbb{S}^{d-1})$ is dense in $|D_s|^{-1/2}L^2(\widetilde{C})$, it suffices, for $w \in C_0^\infty(\mathbb{R}\times\mathbb{S}^{d-1})$, to find a sequence $\phi_n \in \mathrm{Sol}_{\mathrm{sc}}(P_0)$ such that $\rho\phi_n \to w$ in $|D_s|^{-1/2}L^2(\widetilde{C})$.

We will use freely the notation introduced in Section 8D. We first fix a Cauchy surface Σ in (M, g) as in Section 8D2 to the future of $\operatorname{supp} w$. Note that, since w vanishes near $s = -\infty$, we know that w belongs to the space $\widetilde{H}^1_0(\widetilde{C}_0)$ introduced in Proposition 8.8. By Theorem 8.7 and Proposition 8.8, there exists f in the energy space $\mathscr{E}_0(\Sigma_0)$ such that $w = R\circ Tf$. Since $C_0^\infty(\Sigma_0)\oplus C_0^\infty(\Sigma_0)$ is dense in $\mathscr{E}_0(\Sigma_0)$, there exists a sequence $f_n \in C_0^\infty(\Sigma_0)\oplus C_0^\infty(\Sigma_0)$ such that $f_n \to f$ in $\mathscr{E}_0(\Sigma_0)$. By Theorem 8.7 and Proposition 8.8 we have $R\circ Tf_n \to w$ in $\widetilde{H}^1_0(\widetilde{C}_0)$, hence also $R\circ Tf_n \to w$ in $|D_s|^{-1/2}L^2(\widetilde{C})$, by Remark 8.9.

Let $\phi_n \in \mathrm{Sol}_{\mathrm{sc}}(P_0)$ be the solution with Cauchy data f_n on Σ_0. Then $\rho\phi_n = R\circ Tf_n \to w$ in $|D_s|^{-1/2}L^2(\widetilde{C})$, which completes the proof of the theorem. □

8D. *A characteristic Cauchy problem in M_0.* From Lemma 8.1 we see that, to deduce purity of the bulk state from the purity of the boundary state, the range of ρ in $\mathscr{H}(\widetilde{C})$ should be sufficiently large. One way to ensure this is to solve a *characteristic Cauchy problem* in M_0, that is, to construct an inverse for ρ. If M has a *compact* Cauchy surface, the characteristic problem was shown to be well posed in *energy spaces* by Hörmander [1990b]. With some care those results can be used in our situation.

8D1. *Characteristic Cauchy problem for compact Cauchy surfaces.* We recall an important result of [Hörmander 1990b] on the characteristic Cauchy problem in energy spaces, whose framework is as follows:

One considers a spacetime $(\widetilde{M}, \tilde{g})$ for $\widetilde{M} = \mathbb{R} \times \widetilde{\Sigma}$, where $\widetilde{\Sigma}$ is a smooth *compact* manifold and $\tilde{g} = -\tilde{\beta}(t, \mathrm{x})\, dt^2 + \tilde{h}_{ij}(t, \mathrm{x})\, d\mathrm{x}^i\, d\mathrm{x}^j$. One also fixes a real function $\tilde{r} \in C^\infty(\widetilde{M})$.

If $\widetilde{\Sigma}_1$ is a Cauchy hypersurface in $(\widetilde{M}, \tilde{g})$, we will denote by

$$\widetilde{U}_{\widetilde{\Sigma}_1} : C^\infty(\widetilde{\Sigma}_1) \oplus C^\infty(\widetilde{\Sigma}_1) \to C^\infty(\widetilde{M})$$

the Cauchy evolution operator for $-\Box_{\tilde{g}} + \tilde{r}$, so that $\phi = \widetilde{U}_{\widetilde{\Sigma}_1} f$ solves

$$\begin{cases} -\Box_{\tilde{g}}\phi + \tilde{r}\phi = 0, \\ \phi\restriction_{\widetilde{\Sigma}_1} = f^0, \\ n^\mu \nabla_\mu \phi\restriction_{\widetilde{\Sigma}_1} = f^1. \end{cases}$$

A hypersurface $\widetilde{C}$ of the form

$$\widetilde{C} = \{(F(\mathrm{x}), \mathrm{x}) : \mathrm{x} \in \widetilde{\Sigma}\}, \quad F \text{ Lipschitz}, \tag{8-7}$$

is called *spacelike* (resp. *weakly spacelike*) if

$$\sup_{\mathrm{x}\in\widetilde{\Sigma}} \big(-\beta^{-1}(F(\mathrm{x}), \mathrm{x}) + \partial_i F(\mathrm{x}) h^{ij}(F(\mathrm{x}), \mathrm{x}) \partial_j F(\mathrm{x})\big) < 0 \quad (\text{resp. } \le 0).$$

If F is smooth then of course $\widetilde{C}$ is spacelike (resp. weakly spacelike) if and only if all tangent vectors at each point of $\widetilde{C}$ are spacelike (resp. spacelike or null).

Since $\widetilde{\Sigma}$ is compact and F Lipschitz, the Sobolev space $H^1(\widetilde{C})$ and of course $L^2(\widetilde{C})$ are well defined, for example by identifying $\widetilde{C}$ with $\widetilde{\Sigma}$ and using the Riemannian metric $\tilde{h}_{ij}(0, \mathrm{x})\, d\mathrm{x}^i\, d\mathrm{x}^j$ on $\widetilde{\Sigma}$ to equip $\widetilde{C}$ with a density $d\nu_{\widetilde{C}}$.

One also needs the measure

$$d\nu^0_{\widetilde{C}} = (\beta^{-1} - h^{ij}\, \partial_i \widetilde{F}\, \partial_j \widetilde{F})\, d\nu_{\widetilde{C}},$$

which vanishes if $\widetilde{C}$ is a null hypersurface.

We now set

$$\mathscr{E}(\widetilde{C}) := H^1(\widetilde{C}) \oplus L^2(\widetilde{C}, d\nu^0_{\widetilde{C}}). \tag{8-8}$$

Note that if $\widetilde{C}$ is spacelike (i.e., a Cauchy hypersurface), then $\mathscr{E}(\widetilde{C}) = H^1(\widetilde{C}) \oplus L^2(\widetilde{C})$.

Theorem 8.5 [Hörmander 1990b]. *Let $\widetilde{\Sigma}_1$ be any Cauchy hypersurface in $\widetilde{M}$ and let $\widetilde{C}$ be weakly spacelike of the form* (8-7). *Then the map*

$$\widetilde{T} : \mathscr{E}(\widetilde{\Sigma}_1) \to \mathscr{E}(\widetilde{C}), \quad f \mapsto ((\widetilde{U}_{\widetilde{\Sigma}_1} f)\restriction_{\widetilde{C}}, (\beta^{-1}\partial_t \widetilde{U}_{\widetilde{\Sigma}_1} f)\restriction_{\widetilde{C}}),$$

is a homeomorphism.

Note that, if $\widetilde{C}$ is characteristic, then $L^2(\widetilde{C}, d\nu^0_{\widetilde{C}}) = \{0\}$ and $\mathscr{E}(\widetilde{C}) = H^1(\widetilde{C})$, so one obtains as a particular case the solvability of the characteristic Cauchy problem in energy spaces.

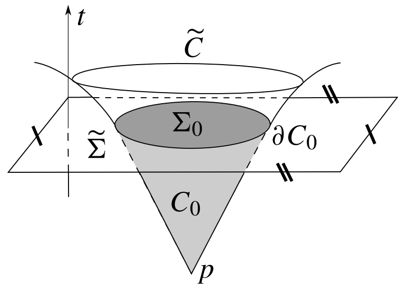

Figure 2. The modified cone $\widetilde{C}$.

8D2. *Embedding M_0 into $\widetilde{M}$.* We will use Hörmander's result, recalled above, to solve a characteristic Cauchy problem in M_0 in an arbitrary neighborhood of p. The first task is to locally embed M into a spacetime $\widetilde{M}$ as above.

We fix a Cauchy hypersurface Σ to the future of p and identify M with $\mathbb{R}\times\Sigma$ equipped with

$$g = -\beta(t,\mathrm{x})\,dt^2 + h_{ij}(t,\mathrm{x})\,d\mathrm{x}^i\,d\mathrm{x}^j.$$

We set $\Sigma_0 = \Sigma\cap M_0$ and fix an open, precompact set U such that $J^-(\Sigma_0)\cap J^+(p)\subset U$.

The following lemma shows that, over U, C can be parametrized by Σ.

Lemma 8.6. *There exists a bounded, Lipschitz function F defined on Σ such that*

$$\overline{C}\cap U = \{(t,\mathrm{x}) : t = F(\mathrm{x})\}\cap U.$$

Proof. The proof is given in Appendix A6. □

We next embed Σ_0 into a smooth compact manifold $\widetilde{\Sigma}$. We consider the spacetime $\widetilde{M} = \mathbb{R}\times\widetilde{\Sigma}$ and extend F to a Lipschitz function $\widetilde{F}$ on $\widetilde{\Sigma}$ and g to a metric $\tilde{g}$ as in Section 8D1. We set

$$\widetilde{C} = \{t = \widetilde{F}(\mathrm{x})\}\subset\widetilde{M}$$

and define

$$C_0 := (J^-(\Sigma_0; M)\cap C)\cup\{p\}. \tag{8-9}$$

C_0 is an open subset of $\overline{C}$, with $\overline{C}_0$ compact in M and

$$\partial\Sigma_0 = \partial C_0. \tag{8-10}$$

We claim that we can choose the embedding $\Sigma_0\subset\widetilde{\Sigma}$ and the extensions $\widetilde{F}$ and $\tilde{g}$ so that

$$J^-(\widetilde{\Sigma}\backslash\overline{\Sigma}_0; \widetilde{M})\cap\overline{C}_0 = \varnothing, \tag{8-11}$$

$$\widetilde{C}\ \text{ is weakly spacelike in }\ \widetilde{M}. \tag{8-12}$$

This is clearly possible by modifying Σ, F and g only outside a large open set U and using that the embedding of (M_0, g) into (M, g) is causally compatible; see (2-3).

The situation is summarized in Figure 2. Identification symbols (a single and double bar) are used to stress that $\widetilde{\Sigma}$ is compact.

8D3. *Sobolev spaces.* We now recall some well-known facts about Sobolev spaces. If Ω is a relatively compact open set in a compact manifold X with smooth boundary $\partial\Omega$, then $H^1_0(\Omega)$ — defined as the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$ — can also be characterized as $H^1_0(\Omega)=\{u\in H^1(\Omega): u\restriction_{\partial\Omega}=0\}$. The restriction operator $r_\Omega: H^1(X)\to H^1(\Omega)$ is surjective from $H^1_{\partial\Omega}(X)=\{u\in H^1(X): u\restriction_{\partial\Omega}=0\}$ to $H^1_0(\Omega)$, with right inverse $e_\Omega: H^1_0(\Omega)\to H^1_{\partial\Omega}(X)$ equal to the extension by 0 in $X\backslash\Omega$.

We set $\mathscr{E}_0(\Omega):=H^1_0(\Omega)\oplus L^2(\Omega)$ and $\mathscr{E}_{\partial\Omega}(X)=H^1_{\partial\Omega}(X)\oplus L^2(X)$. We will still denote the operator $r_\Omega\oplus r_\Omega:\mathscr{E}_{\partial\Omega}(X)\to\mathscr{E}_0(\Omega)$ by r_Ω and $e_\Omega\oplus e_\Omega:\mathscr{E}_0(\Omega)\to\mathscr{E}_{\partial\Omega}(X)$ by e_Ω.

We will use these facts for $\Omega=\Sigma_0$, C_0 and $X=\widetilde{\Sigma}$, $\widetilde{C}$. If $\Omega=C_0$, then we use the notation in (8-8), i.e., $\mathscr{E}_0(C_0)=H^1_0(C_0)\oplus\{0\}\sim H^1_0(C_0)$, since C_0 is characteristic.

8D4. *Characteristic Cauchy problem.* In the theorem below, we denote by U_{Σ_0} the operator $U_{\widetilde{\Sigma}}\circ e_{\Sigma_0}$, that is, the Cauchy evolution operator (in $\widetilde{M}$) for Cauchy data in $\mathscr{E}_0(\Sigma_0)$ (extended by 0 in $\widetilde{\Sigma}\backslash\Sigma_0$).

Theorem 8.7. *The map*

$$T:\mathscr{E}_0(\Sigma_0)\to\mathscr{E}_0(C_0),\qquad f\mapsto (U_{\Sigma_0}f)\restriction_{C_0},$$

is a homeomorphism.

Proof. We will prove the theorem by reducing ourselves to Theorem 8.5. We first claim that

$$T=r_{C_0}\circ\widetilde{T}\circ e_{\Sigma_0}. \tag{8-13}$$

In fact this follows from the fact that $e_{\Sigma_0}:\mathscr{E}_0(\Sigma_0)\to\mathscr{E}(\widetilde{\Sigma})$ is the extension by 0.

By Theorem 8.5, this implies that $T:\mathscr{E}_0(\Sigma_0)\to\mathscr{E}(C_0)$. Moreover, by finite speed of propagation, if $f\in C_0^\infty(\Sigma_0)\oplus C_0^\infty(\Sigma_0)$ then Tf vanishes near ∂C_0, hence T maps continuously $\mathscr{E}_0(\Sigma_0)$ into $\mathscr{E}_0(C_0)$.

We next claim that $S=r_{\Sigma_0}\circ\widetilde{T}^{-1}\circ e_{C_0}$ is a right inverse to T. In fact, let $g\in\mathscr{E}_0(C_0)$ and $\tilde f=\widetilde{T}^{-1}\circ e_{C_0}g=(\tilde f^0,\tilde f^1)\in\mathscr{E}(\widetilde{\Sigma})$. Since $\partial\Sigma_0=\partial C_0$, we have $\tilde f^0\restriction_{\partial\Sigma_0}=g\restriction_{\partial C_0}=0$, hence $e_{\Sigma_0}\circ r_{\Sigma_0}\tilde f\in\mathscr{E}(\widetilde{\Sigma})$. Since $\tilde f-e_{\Sigma_0}\circ r_{\Sigma_0}\tilde f$ vanishes on $\overline{\Sigma}_0$, we obtain by (8-11) and finite speed of propagation that

$$r_{C_0}\circ\widetilde{T}(\tilde f-e_{\Sigma_0}\circ r_{\Sigma_0}\tilde f)=0,$$

hence $T\circ Sg=r_{C_0}\circ\widetilde{T}\tilde f=r_{C_0}\circ e_{C_0}g=g$. This completes the proof of the theorem. □

8E. *Sobolev space on the cone in null coordinates.* Let us set

$$R: C^\infty(C)\to C^\infty(\mathbb{R}\times\mathbb{S}^{d-1}),\qquad g\mapsto\beta^{-1}g(s,\theta).$$

The goal in this subsection is to describe more precisely the image of $H^1_0(C_0)$ under R.

We will denote by $\widetilde{C}_0\subset\mathbb{R}\times\mathbb{S}^{d-1}$ the image of C_0 under the map $q\mapsto(s(q),\theta(q))$ for $q\in C$, where the coordinates (s,θ) are as constructed in Lemma 2.6. Using that $\partial C_0=\partial\Sigma_0$ is spacelike and included in C, we easily obtain from Lemma 2.7 that $\widetilde{C}_0$ is of the form

$$\widetilde{C}_0=\{(s,\theta)\in\mathbb{R}\times\mathbb{S}^{d-1}: s<s_0(\theta)\}$$

for some smooth function s_0. To simplify notation, the measure $|m|^{1/2}(\theta)\,d\theta$ on $\mathbb{S}^{d-1}$ will be simply denoted by $d\theta$. We also set $r=\mathrm{e}^s$.

Proposition 8.8. *Assume $d = \dim M - 1 \geq 3$. Then the image of $H_0^1(C_0)$ under R equals the completion of $C_0^\infty(\widetilde{C}_0)$ under the norm*

$$\|\psi\|_1 := \left(\int_{\widetilde{C}_0} (r^{-1}|\partial_s\psi|^2 + r^{-1}|\partial_\theta\psi|^2 + r^{-1}|\psi|^2)\, ds\, d\theta\right)^{\frac{1}{2}}.$$

We will denote this space by $\widetilde{H}_0^1(\widetilde{C}_0)$.

Remark 8.9. Since $r \leq r_0$ on C_0, we see that $\widetilde{H}_0^1(\widetilde{C}_0)$ injects continuously into $|D_s|^{-1/2}L^2(\mathbb{R}\times\mathbb{S}^{d-1})$.

Proof. We recall that (v, ψ) (see (2-4)) are coordinates on C such that the topology in $H_0^1(C_0)$ is given by the norm

$$\left(\int_{C_0} (|v|^{d-1}|\partial_v g|^2 + |v|^{d-3}|\partial_\psi g|^2 + |v|^{d-1}|g|^2)\, dv\, d\psi\right)^{\frac{1}{2}}.$$

Recall that we have set $r = \mathrm{e}^s$. A function $g \in H_0^1(C_0)$ expressed in the coordinates (s, θ) or (r, θ) will still be denoted by g. Similarly, the image of $\widetilde{C}_0$ under the map $(s, \theta) \mapsto (\mathrm{e}^s, \theta)$ will still be denoted by $\widetilde{C}_0$.

From Lemma 2.6(3) and a routine computation, we see that an equivalent norm on $H_0^1(C_0)$ is

$$\left(\int_{\widetilde{C}_0} (r^{d-1}|\partial_r g|^2 + r^{d-3}|\partial_\theta g|^2 + r^{d-1}|g|^2\, dr\, d\theta)\right)^{\frac{1}{2}}. \tag{8-14}$$

Since $d = \dim M - 1 \geq 3$, Hardy's inequality $-\Delta \geq C|x|^{-2}$ holds on $L^2(\mathbb{R}^d)$. Considering (r, θ) as polar coordinates on $\mathbb{R}^d$, we obtain that

$$\int_{\widetilde{C}_0} r^{d-1}|\partial_r g|^2 + r^{d-3}|\partial_\theta g|^2\, dr\, d\theta \geq C\int_{\widetilde{C}_0} r^{d-3}|g|^2\, dr\, d\theta, \quad g \in H_0^1(C_0).$$

Therefore, adding a term $r^{d-3}|g|^2$ under the integral in (8-14) yields an equivalent norm on $H_0^1(C_0)$. Since r is bounded on $\widetilde{C}_0$, this term dominates the term $r^{d-1}|g|^2$ and we finally obtain that the topology of $H^1(C_0)$ is given by the norm

$$\left(\int_{\widetilde{C}_0} (r^{d-1}|\partial_r g|^2 + r^{d-3}|\partial_\theta g|^2 + \alpha r^{d-3}|g|^2)\, dr\, d\theta\right)^{\frac{1}{2}},$$

where the constant $\alpha > 0$ can be chosen arbitrarily large. Going back to coordinates (s, θ), we obtain the norm

$$\left(\int_{\widetilde{C}_0} (r^{d-2}|\partial_s g|^2 + r^{d-2}|\partial_\theta g|^2 + \alpha r^{d-2}|g|^2)\, ds\, d\theta\right)^{\frac{1}{2}}. \tag{8-15}$$

For two functions $m, n \in C^\infty(\mathbb{C}_0)$ we write $m \sim n$ if $m = r_0 n$ for some $r_0, r_0^{-1} \in S^0$, where the class S^0 is as defined in Section 2E. We have $\beta \sim r^{-(d-1)/2}$, hence

$$\partial_s\beta,\, \partial_\theta\beta \sim r^{-(d-1)/2}. \tag{8-16}$$

Setting $\psi = Rg = \beta^{-1}g$, we have

$$\partial_s g = \beta\,\partial_s\psi + (\partial_s\beta)\psi \quad \text{and} \quad \partial_\theta g = \beta\,\partial_\theta\psi + (\partial_\theta\beta)\psi.$$

Then, using (8-16) and choosing $\alpha \gg 1$ in (8-15), we obtain that (8-15) is equivalent to

$$\left(\int_{\widetilde{C}_0} (r^{-1}|\partial_s \psi|^2 + r^{-1}|\partial_\theta \psi|^2 + r^{-1}|\psi|^2)\, ds\, d\theta\right)^{\frac{1}{2}}. \tag{8-17}$$

This completes the proof of the proposition. □

9. Change of null coordinates

The map $\rho : \mathrm{Sol}_{\mathrm{sc}}(P_0) \to \mathcal{H}(\widetilde{C})$ introduced in Definition 4.1 depends on the choice of the null coordinates (s, θ) on C, i.e., on the choice of the initial hypersurface S used in Lemma 2.6 to construct (s, θ). In this section we discuss how our class of Hadamard states depends on the above choice.

9A. *New null coordinates.* We fix a reference hypersurface S in C, yielding null coordinates (s, θ) near C such that $g\restriction_C$ is given by (2-6) and $S = \{f = s = 0\}$.

We choose another hypersurface $\tilde{S}$ transverse to $\nabla^a f$ in C, hence

$$\tilde{S} = \{f = 0,\ s = b(\theta)\} \quad \text{for some } b \in C^\infty(\mathbb{S}^{d-1}). \tag{9-1}$$

Since $\nabla^a f\restriction_C = \partial_s$, we obtain that the new coordinates $(\tilde{s}, \tilde{\theta})$ obtained from Lemma 2.6 with S replaced by $\tilde{S}$ are given by

$$\tilde{\theta} = \theta, \quad \tilde{s}(s, \theta) = s - b(\theta). \tag{9-2}$$

We then have

$$g\restriction_C = -2\, df\, d\tilde{s} + \tilde{h}_{ij}(\tilde{s}, \theta)\, d\theta^i\, d\theta^j$$

and a standard computation shows that $|h|(\tilde{s}, \theta) = |h|(s, \theta)$, hence $\tilde{\beta}(\tilde{s}, \theta) = \beta(s, \theta)$. Denoting by $\tilde{\rho}$ the analog of ρ in Definition 4.1 for the new coordinates $(\tilde{s}, \theta)$ we then have

$$\tilde{\rho}\phi = U\rho\phi, \quad \phi \in \mathrm{Sol}_{\mathrm{sc}}(P_0), \tag{9-3}$$

where

$$U : \mathcal{H}(\widetilde{C}) \to \mathcal{H}(\widetilde{C}), \quad g \mapsto Ug(s, \theta) = g(s + b(\theta), \theta).$$

The map U is symplectic on $(\mathcal{H}(\widetilde{C}), \sigma_C)$ and unitary on $L^2(\widetilde{C})$ with $U^* D_s U = D_s$.

Proposition 9.1. *If $A \in \widetilde{\Psi}^{-\infty,p}(\widetilde{C})$ then $UAU^{-1} \in \widetilde{\Psi}^{-\infty,p}(\widetilde{C})$.*

Remark 9.2. The above invariance property does not hold for the classes $\Psi^{m,p}(\widetilde{C})$ since, for example, the classes $\Psi^{m,p}(\mathbb{R} \times \mathbb{R}^{d-1})$ are not even preserved by linear changes of variables $(s, y) \mapsto (s + Ay, y)$.

Proof. We will use the Beals criterion explained in Section 6C, which implies that $B \in \widetilde{\Psi}^{-\infty,p}(\widetilde{C})$ if and only if, for any functions $g_1, \dots g_n \in C^\infty(\mathbb{S}^{d-1})$ and smooth vector fields $X_1, \dots, X_m$ on $\mathbb{S}^{d-1}$ and any $N \in \mathbb{N},\ k, k' \in \mathbb{R}$, one has

$$\mathrm{ad}_{X_1} \cdots \mathrm{ad}_{X_m}\, \mathrm{ad}_{g_1} \cdots \mathrm{ad}_{g_n}\, B : H^{k,k'}(\widetilde{C}) \to H^{k+N,k'-p+n}(\widetilde{C}). \tag{9-4}$$

To simplify notation, we rewrite (9-4) as

$$\mathrm{ad}_{\overline{X}}^{\alpha}\, \mathrm{ad}_{\overline{g}}^{\beta}\, B : H^{k,k'}(\widetilde{C}) \to H^{k+N,k'+p+|\beta|}(\widetilde{C}), \tag{9-5}$$

denoting by $\bar{X}$ and $\bar{g}$ an arbitrary n-tuple of vector fields and m-tuple of functions, respectively.

If g is a function on $\mathbb{S}^{d-1}$, considered as a multiplication operator, and X is a vector field on $\mathbb{S}^{d-1}$, we have

$$U^{-1}gU = g, \quad U^{-1}XU = X + (X \cdot db)\partial_s, \quad U^{-1}\partial_s U = \partial_s. \tag{9-6}$$

Now let $A \in \widetilde{\Psi}^{-\infty,p}(\widetilde{C})$. For $\psi \in C^\infty(\mathbb{S}^{d-1}\times\mathbb{S}^{d-1})$, let us denote by A_ψ the operator with distributional kernel $A(s_1, s_2, \theta_1, \theta_2)\psi(\theta_1, \theta_2)$. By the well-known properties of the pseudodifferential calculus on $\mathbb{S}^{d-1}$, we know that if $\psi = 1$ in some neighborhood of the diagonal then $A - A_\psi \in \widetilde{\Psi}^{-\infty,-\infty}(\widetilde{C})$ or, equivalently, maps $H^{k,k'}(\widetilde{C})$ into $H^{k+N,k'+N}(\widetilde{C})$ for any k, k' and N. Using (9-6) this implies that $U(A - A_\psi)U^{-1}$ has the same property, hence belongs to $\widetilde{\Psi}^{-\infty,-\infty}(\widetilde{C})$.

Therefore we can replace A by A_ψ and assume that the kernel of A is supported in $\mathbb{R}\times\mathbb{R}\times\Omega$, where Ω is an arbitrarily small neighborhood of the diagonal in $\mathbb{S}^{d-1}\times\mathbb{S}^{d-1}$. Introducing a smooth partition of unity $1 = \sum_1^M \chi_i$ on $\mathbb{S}^{d-1}$, we see that we can replace A by $\chi A \chi$, where $\chi \in C^\infty(\mathbb{S}^{d-1})$ is supported in a small neighborhood of a point $\theta_0 \in \mathbb{S}^{d-1}$. We pick local coordinates $\theta_1, \dots, \theta_{d-1}$ near θ_0 and rewrite (9-5) as

$$\langle\partial_s\rangle^{k+N}\langle\partial_\theta\rangle^{k'-p+|\beta|}\,\mathrm{ad}^\alpha_{\bar{X}}\,\mathrm{ad}^\beta_{\bar{g}}\,A\langle\partial_s\rangle^{-k}\langle\partial_\theta\rangle^{-k'} \in B(L^2(\widetilde{C})). \tag{9-7}$$

We now set $A' = UAU^{-1}$. Note first that if the kernel of A is supported in $\mathbb{R}\times\mathbb{R}\times\Omega$ then so is the kernel of A', hence by the above discussion it suffices to check that A' satisfies (9-7). Let us set $U^{-1}XU = X'$ if X is a vector field on $\mathbb{S}^{d-1}$ and, in particular, $\partial'_\theta = U^{-1}\partial_\theta U = \partial_\theta + \partial_\theta b\partial_s$. Then an easy computation yields

$$\begin{aligned}\langle\partial_s\rangle^{k+N}\langle\partial_\theta\rangle^{k'-p+|\beta|}\,\mathrm{ad}^\alpha_{\bar{X}}\,\mathrm{ad}^\beta_{\bar{g}}\,UAU^{-1}\langle\partial_s\rangle^{-k}\langle\partial_\theta\rangle^{-k'}&\\ = U\langle\partial_s\rangle^{k+N}\langle\partial'_\theta\rangle^{k'-p+|\beta|}\,\mathrm{ad}^\alpha_{\bar{X}'}\,\mathrm{ad}^\beta_{\bar{g}}\,A\langle\partial_s\rangle^{-k}\langle\partial'_\theta\rangle^{-k'}U^{-1}.&\end{aligned} \tag{9-8}$$

Using (9-6) and the fact that $A \in \widetilde{\Psi}^{-\infty,p}(\widetilde{C})$, we obtain that

$$\mathrm{ad}^\alpha_{\bar{X}'}\,\mathrm{ad}^\beta_{\bar{g}}\,A \in \widetilde{\Psi}^{-\infty,p-|\beta|}(\widetilde{C}) \quad \text{and} \quad \langle\partial_s\rangle^{N}\langle\partial_\theta\rangle^{k'-p+|\beta|}\,\mathrm{ad}^\alpha_{\bar{X}'}\,\mathrm{ad}^\beta_{\bar{g}}\,A\langle\partial_s\rangle^{N}\langle\partial_\theta\rangle^{-k'} \in B(L^2(\widetilde{C}))$$

for any $N \in \mathbb{N}$. It follows that the left-hand side of (9-8) belongs to $B(L^2(\widetilde{C}))$ if, for any $s \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that

$$\langle\partial_s\rangle^{-N}\langle\partial'_\theta\rangle^{s}\langle\partial_\theta\rangle^{-s},\ \langle\partial_s\rangle^{-N}\langle\partial_\theta\rangle^{s}\langle\partial'_\theta\rangle^{-s} \in B(L^2(\widetilde{C})). \tag{9-9}$$

Let us now prove (9-9). The first statement of (9-9) is easy to check for $s \in \mathbb{N}$, using that $\partial'_\theta = \partial_\theta + \partial_\theta b\,\partial_s$. Conjugation by U gives the second statement for $s \in \mathbb{N}$. By duality and interpolation, we then obtain (9-9) for arbitrary s, which completes the proof of the proposition. □

From Proposition 9.1 and the fact that $U^* D_s U = D_s$, we immediately obtain the following result:

Proposition 9.3. *The classes of Hadamard states obtained in Theorems 7.4 and 8.2 are independent of the choice of the null coordinates* (s, θ).

Appendix

A1. ***Stokes formula.*** Let (M, g) an orientable, oriented pseudo-Riemannian manifold of dimension n. We denote by $d\,\mathrm{Vol}_g \in \bigwedge^n(M)$ the associated volume form and by $d\mu_g = |d\,\mathrm{Vol}_g|$ the associated density.

Let $\Sigma \subset M$ a smooth submanifold of codimension 1 and $\iota : \Sigma \to M$ the natural injection, which induces $\iota^* : \bigwedge(M) \to \bigwedge(\Sigma)$. From the orientation of M and a continuous transverse vector field $v \in T_\Sigma M$, we obtain an induced orientation of Σ. If $\Sigma \subset \partial U$ for an open set $U \subset M$ with piecewise smooth boundary ∂U, we choose v pointing outwards.

If $\omega \in \bigwedge^n(M)$ and $X \in TM$, then $X \lrcorner\, \omega \in \bigwedge^{n-1}(M)$ and one sets

$$\iota_X^* \omega := \iota^*(X \lrcorner\, \omega) \in \textstyle\bigwedge^{n-1}(\Sigma).$$

Similarly, if $\mu = |\omega|$ is a density on M, we set $\iota_X^* \mu := |\iota_X \omega|$, which is a density on Σ.

If ∇_a is the Levi-Civita connection associated to g then

$$\nabla_a X^a \, d\,\mathrm{Vol}_g = d(X \lrcorner\, d\,\mathrm{Vol}_g),$$

which, applying Stokes formula

$$\int_U d\omega = \int_{\partial U} \iota^* \omega, \quad \omega \in \textstyle\bigwedge^{n-1}(M), \tag{A-1}$$

to $\omega = \iota_X^* \, d\,\mathrm{Vol}_g$ yields

$$\int_U \nabla_a X^a \, d\,\mathrm{Vol}_g = \int_{\partial U} \iota_X^* \, d\,\mathrm{Vol}_g\,. \tag{A-2}$$

Noncharacteristic boundaries. Assume first $\Sigma \subset \partial U$ is *noncharacteristic*, that is, the one-dimensional space

$$T_x(\Sigma)^{\mathrm{ann}} \subset T_x M^*$$

is not *null* (the superscript "ann" denotes the annihilator). It follows that the metric $h := \iota^* g$ on Σ is nondegenerate (in the Lorentzian case, one typically assume that Σ is spacelike; then $h = \iota^* g$ is Riemannian). Let $n \in T_\Sigma M$ be the unit, outward-pointing normal vector field to σ. Then

$$d\,\mathrm{Vol}_h = \iota_n^* \, d\,\mathrm{Vol}_g \quad \text{and} \quad \iota_X^* \, d\,\mathrm{Vol}_g = X^a n_a \, d\,\mathrm{Vol}_h, \tag{A-3}$$

hence

$$\int_\Sigma \iota_X^* \, d\,\mathrm{Vol}_g := \int_\Sigma X^a n_a \, d\sigma_h.$$

If all of ∂U is noncharacteristic, then from (A-2) we obtain Gauss's formula

$$\int_U \nabla_a X^a \, d\mu_g = \int_\Sigma X^a n_a \, d\sigma_h, \tag{A-4}$$

where $d\sigma_h = |d\,\mathrm{Vol}_h|$.

Characteristic boundaries. Assume now that Σ is characteristic. Then there is no normal vector field anymore. To express the right-hand side of (A-2), one chooses a defining function f for Σ, i.e., such that $f=0$ and $df\neq 0$ on Σ, and completes f with coordinates $y^1,\dots,y^{n-1}$ such that $df\wedge dy^1\wedge\cdots\wedge dy^{n-1}$ is positively oriented. Then, computing in the coordinates $f, y^1,\dots,y^{n-1}$, one sees that

$$\iota_X^* \, d\,\mathrm{Vol}_g = X^a\nabla_a f|g|^{1/2}\,dy^1\wedge\cdots\wedge dy^{n-1},$$

hence

$$\int_\Sigma \iota_X^*\, d\,\mathrm{Vol}_g = \int_\Sigma X_a\nabla^a f|g|^{1/2}\,dy^1\wedge\cdots\wedge dy^{n-1} \tag{A-5}$$

In the general case we can, for example, split ∂U as $\Sigma_1\cup\Sigma_2$, where Σ_1 is noncharacteristic and Σ_2 is characteristic, and obtain

$$\int_U \nabla_a X^a\,d\mu_g = \int_{\Sigma_1} X^a n_a\,d\sigma_h + \int_{\Sigma_2} X_a\nabla^a f|g|^{1/2}\,dy^1\wedge\cdots\wedge dy^{n-1}. \tag{A-6}$$

A2. *Conformal transformations.* In this section we briefly discuss conformal transformations of a globally hyperbolic spacetime (M, g). Let $\omega\in C^\infty(M)$ be strictly positive and consider the conformally related metric

$$g' = \omega^2 g.$$

Set

$$P = -\nabla^a\nabla_a + \frac{n-2}{4(n-1)}R,$$

where R is the scalar curvature. For this special choice of the lower-order terms, the conformal transformation $g\to g'$ amounts to

$$P' = \omega^{-n/2-1}P\omega^{n/2-1}.$$

This entails that the causal propagators are related by $E' = \omega^{-n/2+1}E\omega^{n/2+1}$. One concludes that multiplication by $\omega^{-n/2+1}$ induces a symplectic map

$$(\mathrm{Sol}_{\mathrm{sc}}(P),\sigma) \xrightarrow{\omega^{-n/2+1}} (\mathrm{Sol}_{\mathrm{sc}}(P'),\sigma'), \tag{A-7}$$

where σ and σ' are defined as in (3-2) using the respective volume densities.

We apply this discussion to (M_0, g) and the conformally related spacetime with metric $g'=\omega^2 g$. In the setting of Section 4A, there is a monomorphism of symplectic spaces

$$(\mathrm{Sol}_{\mathrm{sc}}(P_0),\sigma_0) \xrightarrow{\rho} (\mathscr{H}(\widetilde{C}),\sigma_C).$$

By (A-7) we also have a monomorphism

$$(\mathrm{Sol}_{\mathrm{sc}}(P_0'),\sigma_0') \xrightarrow{\rho\circ\omega^{n/2-1}} (\mathscr{H}(\widetilde{C}),\sigma_C).$$

Therefore, one can construct states for the conformally related spacetime using the bulk-to-boundary correspondence with a modified trace map $\rho' = \rho\circ\omega^{n/2-1}$.

A3. ***Proof of Lemma 2.7.*** We fix a point $q \in C$ and complete the coordinate $x^0 = f$ by local coordinates $\bar{x} = (x^1, \dots, x^d)$ near q. The functions s and θ_k defined on C are denoted by $s(\bar{x})$ and $\theta_k(\bar{x})$, since $\bar{x}$ are local coordinates on C. We denote by $h(\bar{x})$ the restriction of g^{-1} to T^*C. Note that the fact that C is null implies that $g^{00}(0, \bar{x}) \equiv 0$ and that from Lemma 2.6 we have

$$g^{i0}(\bar{x})\,\partial_i s(\bar{x}) = -1, \quad g^{i0}(\bar{x})\,\partial_i \theta_k(\bar{x}) = 0. \tag{A-8}$$

If X is a null vector, orthogonal to $C \cap \{s(\bar{x}) = s(q)\}$ and transverse to C, we obtain that

$$gX = \lambda\big(\tfrac{1}{2}\nabla_i s \nabla^i s, \nabla_i s\big), \quad \lambda \in \mathbb{R}.$$

Let us denote for the moment by $\tilde{s}$ and $\tilde{\theta}_k$ the extensions of s and θ_k outside C, which are constant along the flow of X. We obtain that, on C,

$$d\tilde{s} = \big(\tfrac{1}{2} ds \cdot h\, ds, ds\big), \quad d\tilde{\theta}_k = (ds \cdot h\, d\theta_k, d\theta_k).$$

Using also $df = (1, 0, \dots, 0)$ and (A-8), a routine computation leads to the following identities on C:

$$\begin{aligned} &df \cdot g^{-1}\, df = d\tilde{s} \cdot g^{-1}\, d\tilde{s} = df \cdot g^{-1}\, d\tilde{\theta}_k = d\tilde{s} \cdot g^{-1}\, d\tilde{\theta}_k = 0, \\ &df \cdot g^{-1}\, d\tilde{s} = d\tilde{s} \cdot g^{-1}\, df = -1, \\ &d\tilde{\theta}_k \cdot g^{-1}\, d\tilde{\theta}_l = \partial_i \theta_k h^{ij} \partial_j \theta_l. \end{aligned}$$

This implies that g is of the form (2-6) on C.

A4. ***Proof of Lemma 2.6.*** Since $(y^0, \bar{y})$ are normal coordinates, we have

$$g\restriction_C = -dv\, dw + \tfrac{1}{4} v^2 m_{ij}(\psi)\, d\psi^i\, d\psi^j + v^2 g_1, \tag{A-9}$$

where $m_{ij}(\psi)\, d\psi^i\, d\psi^j$ is the standard Riemannian metric on $\mathbb{S}^{d-1}$ and g_1 is a smooth pseudo-Riemannian metric in the arguments dv, dw and $v\, d\psi^i$.

We start by expressing f in the normal coordinates $(y^0, \bar{y})$. By Malgrange's preparation theorem [Hörmander 1990a, Theorem 7.5.6] one can write

$$f(y^0, \bar{y}) = m(y^0, \bar{y})((y^0)^2 - |\bar{y}|^2) + a(\bar{y})y^0 + b(\bar{y})$$

for m near $(0, 0)$ and $a, b \in C^\infty$ near 0. Since $C \subset f^{-1}(\{0\})$, we obtain that $b(\bar{y}) = a(\bar{y})|\bar{y}|$ and, since $b \in C^\infty(\mathbb{R}^d)$, necessarily $a \in O(|\bar{y}|^\infty)$. Moreover, from the Hessian of f at p we obtain that $m(0, 0) = 1$.

Going to coordinates (v, w, ψ), we obtain

$$f(v, w, \psi) = m(v, w, \psi)vw + wa(v, w, \psi)$$

for $a \in O(|w - v|^\infty)$. Using also that $m(0, 0, \psi) = 1$, it follows that

$$\partial_v f(v, 0, \psi) = \partial_{\psi^i} f(v, 0, \psi) = 0 \quad \text{and} \quad \partial_w f(v, 0, \psi) = v + r(v, \psi)$$

for $r \in O(|v|^2)$. Using (A-9) to express $(g^{-1})\restriction_C$, we obtain after an easy computation that

$$\nabla^a f = -2v\big((1 + va^0(v, \psi))\,\partial_v + va^i(v, \psi)\,\partial_{\psi^i}\big), \tag{A-10}$$

where a^0 and a^i are smooth, bounded functions near $v=0$.

Let us now prove (1). Using (A-10) we obtain the equation near p

$$(v+v^2a^0(v,\psi))\,\partial_v s+v^2a^i(v,\psi)\,\partial_{\psi^i}s=\tfrac{1}{2}$$

for smooth functions a^0 and a^i. We set $s=\frac{1}{2}\ln(vh(v,\psi))$ and obtain after an elementary computation

$$(1+va^0)\,\partial_v h+a^0h+va^i(v,\psi)\,\partial_{\psi^i}h=0,$$

which we can uniquely solve on $[-\epsilon_1,\epsilon_1]\times\mathbb{S}^{d-1}$ by fixing $h(0,\psi)$. We may fix $h(0,\psi)>0$ to ensure that $s(\epsilon_0,\psi)=0$. We obtain $s=\frac{1}{2}\ln v+\frac{1}{2}\ln h(v,\psi)$ for $h\in C^\infty([-\epsilon_1,\epsilon_1]\times\mathbb{S}^{d-1})$, $h>0$.

It remains to extend s globally to C. To do this it suffices to check that, for any $q\in C$, the integral curve of $\nabla^a f$ through q crosses S at one and only one point. By [Wald 1984, Corollary to Theorem 8.1.2] we know that q can be joined to p by a null geodesic γ. Locally, a null geodesic on C is, modulo reparametrization, an integral curve of $\nabla^a f$. Since $\nabla^a f$ is complete, the whole $\gamma\backslash\{p\}$ is an integral curve of $\nabla^a f$. Hence the integral curve of $\nabla^a f$ through q crosses S. Choosing ϵ_0 in (2-5) small enough, we can ensure that $\nabla^a f\nabla_a v>0$ on S, hence the integral curve through q crosses S at only one point. We can hence extend s globally to C as a C^∞ function.

The proof of (2) is similar. We obtain the equation near p

$$(v+v^2a^0(v,\psi))\,\partial_v\theta^j+v^2a^i(v,\psi)\,\partial_{\psi^i}\theta^j=0$$

or, equivalently,

$$(1+va^0(v,\psi))\,\partial_v\theta^j+va^i(v,\psi)\,\partial_{\psi^i}\theta^j=0,$$

which we can solve in $]-\epsilon_1,\epsilon_1[\times\mathbb{S}^{d-1}$ by imposing $\theta^j(\epsilon_0,\psi)=\psi^j$. The estimate (3) on θ^j is immediate. We extend θ^j to all of C by the same argument as before.

A5. ***Proof of Lemma 6.6.*** We use the characterization of the wavefront set of kernels using oscillatory test functions, which we now recall.

Let $(\tilde{s},\tilde{y})\in C$ and $\lambda\geq 1$. We set, for $(\sigma,\eta)\in\mathbb{R}\times\mathbb{R}^{d-1}$,

$$v_{\sigma,\lambda}(\,\cdot\,)=\chi(\,\cdot\,)\mathrm{e}^{\mathrm{i}\lambda\langle\,\cdot\,,\sigma\rangle}\in C_0^\infty(\mathbb{R})\quad\text{and}\quad w_{\eta,\lambda}(\,\cdot\,)=\psi(\,\cdot\,)\mathrm{e}^{\mathrm{i}\lambda\langle\,\cdot\,,\eta\rangle}\in C^\infty(\mathbb{S}^{d-1}),\tag{A-11}$$

where $\chi\in C_0^\infty(\mathbb{R})$ and $\psi\in C^\infty(\mathbb{S}^{d-1})$ are supported near $\tilde{s}$ and $\tilde{y}$, respectively. We set $u_{(\sigma,\eta),\lambda}=v_{\sigma,\lambda}\otimes w_{\eta,\lambda}$. Note that if V and W are small neighborhoods of $\tilde{\sigma}\in\mathbb{R}$ and $\tilde{\eta}\in\mathbb{R}^{d-1}$, respectively, then for $n_+=\max(n,0)$ we have, uniformly on $U=V\times W$,

$$\|u_{(\sigma,\eta),\lambda}\|_{k,k'}\in\begin{cases}O(\langle\lambda\rangle^{k_++k'_+}),\\ O(\langle\lambda\rangle^{k+k'_+}) & \text{if }\sigma_0\neq 0,\\ O(\langle\lambda\rangle^{k_++k'}) & \text{if }\eta_0\neq 0.\end{cases}\tag{A-12}$$

Let $\widetilde{Y}_1,\widetilde{Y}_2\in T^*C$. Then $(\widetilde{Y}_1,\widetilde{Y}_2)\notin\mathrm{WF}(a)'$ if there exist cutoff functions χ_i and ψ_i with $\chi_i(\tilde{s}_i),\psi_i(\tilde{y}_i)\neq 0$ and neighborhoods $U_i=V_i\times W_i$ of $(\tilde{\sigma}_i,\tilde{\eta}_i)$ such that

$$(u_{(\sigma_1,\eta_1),\lambda}\mid au_{(\sigma_2,\eta_2),\lambda})_{L^2(C)}\in O(\langle\lambda\rangle^{-\infty})\quad\text{uniformly for }(\sigma_i,\eta_i)\in U_i.\tag{A-13}$$

We first prove (1). Let $a \in B^{-\infty}\Psi^{p_2}(C)$ and $\widetilde{Y}_1, \widetilde{Y}_2 \in T^*C$ such that $\tilde{\sigma}_1 \neq 0$ or $\tilde{\sigma}_2 \neq 0$. Then (A-13) follows from (A-12) and the fact that $a : H^{k_1,k_2} \to H^{k_1+m,k_2+p_2}$ for any $m \geq 0$.

We now prove (2). If $a \in \Psi^{p_1,p_2}(C)$ the statement follows from Lemma 6.5. It remains to consider the case $a \in B^{-\infty}\Psi^{p_2}(C)$ and to prove that (A-13) holds if $(\tilde{\sigma}_1, \tilde{\eta}_1) = (0, 0)$ and $(\tilde{\sigma}_2, \tilde{\eta}_2) \neq 0$ or vice versa. If $\tilde{\sigma}_1 \neq 0$ or $\tilde{\sigma}_2 \neq 0$, we have already proved (A-13).

Assume now that $\tilde{\eta}_1 = 0$ and $\tilde{\eta}_2 \neq 0$, the other case being similar. Then we can find cutoff functions $g_i \in C_0^\infty(\mathbb{R}^{d-1})$ supported near $\tilde{\eta}_i$ with disjoint supports such that $(1 - g_i(\lambda^{-1}D_y))u_{(\sigma_i,\eta_i),\lambda} \in O(\lambda^{-\infty})$ in all $H^{k,k'}$ uniformly for $(\sigma_i, \eta_i) \in U$. It follows that

$$(u_{(\sigma_1,\eta_1),\lambda} \mid au_{(\sigma_2,\eta_2),\lambda})_{L^2(C)} = (u_{(\sigma_1,\eta_1),\lambda} \mid g_1(\lambda^{-1}D_y)ag_2(\lambda^{-1}D_y)u_{(\sigma_2,\eta_2),\lambda})_{L^2(C)} + O(\langle\lambda\rangle^{-\infty})$$

uniformly for $(\sigma_i, \eta_i) \in U_i$. By pseudodifferential calculus on $\mathbb{S}^{d-1}$, we know that $g_1(\lambda^{-1}D_y)ag_2(\lambda^{-1}D_y)$ is in $O(\langle\lambda\rangle^{-\infty})$ in $B(H^{k,k'})$ for any $k, k' \in \mathbb{R}$. Combined with (A-12), we obtain (A-13) also if $\tilde{\eta}_1 = 0$ and $\tilde{\eta}_2 \neq 0$. This completes the proof of the lemma.

A6. ***Proof of Lemma 8.6.*** Set $\gamma_{\mathrm{x}} = \{(s, \mathrm{x}) : s \leq 0\}$ for $\mathrm{x} \in \Sigma$. To prove that $\overline{C}$ is the graph of a function F over Σ we have to show that γ_{x} intersects $\overline{C}$ at one and only one point for each $\mathrm{x} \in \Sigma$. Then we have

$$F(\mathrm{x}) = \inf\{s \leq 0 : (s, \mathrm{x}) \in I^+(p)\}.$$

If $F(\mathrm{x}) = -\infty$ then $\gamma_{\mathrm{x}} \subset I^+(p) \cap J^-((0, \mathrm{x})) \subset J^+(p) \cap J^-((0, \mathrm{x}))$. This last set is compact by global hyperbolicity, which is a contradiction. Hence γ_{x} intersects $\overline{C}$. Moreover, if $(t_1, \mathrm{x}) \in \overline{C}$ then $(s, \mathrm{x}) \in J^-(p)$ for all $t_1 \leq s \leq 0$. This shows that γ_{x} intersects $\overline{C}$ at only one point, hence the function F is well defined, and bounded.

Let (T^0, x^0) be the coordinates of p. For $\mathrm{x} \neq \mathrm{x}^0$, C is smooth near $(F(\mathrm{x}), \mathrm{x})$ and ∂_t is transverse to C. By the implicit function theorem this implies that F is smooth near x. Moreover, if $K_1 \subset \Sigma$ is a compact set then dF is uniformly bounded on $K_1 \backslash \{x^0\}$. To prove this it suffices to introduce normal coordinates at p such that, near p, C becomes a neighborhood of the tip of the flat lightcone.

Acknowledgment

The work of Wrochna was partially supported by the FMJH (French Government Program: ANR-10-CAMP-0151-02).

References

[Araki and Shiraishi 1971/72] H. Araki and M. Shiraishi, "On quasifree states of the canonical commutation relations, I", *Publ. Res. Inst. Math. Sci.* **7** (1971/72), 105–120. MR 47 #2387 Zbl 0239.46066

[Bär and Wafo 2015] C. Bär and R. T. Wafo, "Initial value problems for wave equations on manifolds", *Math. Phys. Anal. Geom.* **18**:1 (2015), article ID #7. MR 3316713

[Bär et al. 2007] C. Bär, N. Ginoux, and F. Pfäffle, *Wave equations on Lorentzian manifolds and quantization*, European Mathematical Society, Zürich, 2007. MR 2008j:58041 Zbl 1118.58016

[Benini et al. 2014] M. Benini, C. Dappiaggi, and S. Murro, "Radiative observables for linearized gravity on asymptotically flat spacetimes and their boundary induced states", *J. Math. Phys.* **55**:8 (2014), article ID #082301. MR 3390724 Zbl 1298.83044

[Bony 1997] J.-M. Bony, "Caractérisations des opérateurs pseudo-différentiels", Exposé 23 in *Séminaire sur les équations aux dérivées partielles,* (Palaiseau, 1996–1997), École Polytechnique, Palaiseau, 1997. MR 98m:35233 Zbl 1061.35531

[Borsero and Schulz 2014] M. Borsero and R. Schulz, "Microlocal properties of bisingular operators", *J. Pseudo-Differ. Oper. Appl.* **5**:1 (2014), 43–67. MR 3170600 Zbl 1325.35299

[Brum and Jorás 2015] M. Brum and S. E. Jorás, "Hadamard state in Schwarzschild–de Sitter spacetime", *Classical Quantum Gravity* **32**:1 (2015), article ID #015013. MR 3287021 Zbl 1309.83061

[Cagnac 1981] F. Cagnac, "Problème de Cauchy sur un conoïde caractéristique pour des équations quasi-linéaires", *Ann. Mat. Pura Appl.* (4) **129** (1981), 13–41. MR 84a:35185 Zbl 0486.35023

[Dappiaggi and Siemssen 2013] C. Dappiaggi and D. Siemssen, "Hadamard states for the vector potential on asymptotically flat spacetimes", *Rev. Math. Phys.* **25**:1 (2013), article ID #1350002. MR 3024016 Zbl 1278.81134

[Dappiaggi et al. 2009] C. Dappiaggi, V. Moretti, and N. Pinamonti, "Distinguished quantum states in a class of cosmological spacetimes and their Hadamard property", *J. Math. Phys.* **50**:6 (2009), article ID #062304. MR 2011a:81165 Zbl 1216.81112

[Dappiaggi et al. 2011] C. Dappiaggi, V. Moretti, and N. Pinamonti, "Rigorous construction and Hadamard property of the Unruh state in Schwarzschild spacetime", *Adv. Theor. Math. Phys.* **15**:2 (2011), 355–447. MR 2924234 Zbl 1257.83008

[Dossa 2002] M. Dossa, "Solutions C^∞ d'une classe de problèmes de Cauchy quasi-linéaires hyperboliques du second ordre sur un conoïde caractéristique", *Ann. Fac. Sci. Toulouse Math.* (6) **11**:3 (2002), 351–376. MR 2004i:35224 Zbl 1051.35041

[Fulling et al. 1981] S. A. Fulling, F. J. Narcowich, and R. M. Wald, "Singularity structure of the two-point function in quantum field theory in curved spacetime, II", *Ann. Physics* **136**:2 (1981), 243–272. MR 83e:81057 Zbl 0495.35054

[Gérard and Wrochna 2014] C. Gérard and M. Wrochna, "Construction of Hadamard states by pseudo-differential calculus", *Comm. Math. Phys.* **325**:2 (2014), 713–755. MR 3148100 Zbl 1298.81214

[Häfner and Nicolas 2011] D. Häfner and J.-P. Nicolas, "The characteristic Cauchy problem for Dirac fields on curved backgrounds", *J. Hyperbolic Differ. Equ.* **8**:3 (2011), 437–483. MR 2012i:58021 Zbl 1230.35136

[Hollands and Wald 2015] S. Hollands and R. M. Wald, "Quantum fields in curved spacetime", pp. 513–552 in *General relativity and gravitation: a centennial perspective*, edited by A. Ashtekar et al., Cambridge University Press, 2015.

[Hörmander 1990a] L. Hörmander, *The analysis of linear partial differential operators, I: Distribution theory and Fourier analysis*, 2nd ed., Grundl. Math. Wissen. **256**, Springer, Berlin, 1990. Reprinted in the series Classics in Mathematics, Springer, Berlin, 2003. MR 91m:35001b Zbl 0712.35001

[Hörmander 1990b] L. Hörmander, "A remark on the characteristic Cauchy problem", *J. Funct. Anal.* **93**:2 (1990), 270–277. MR 91m:58154 Zbl 0724.35060

[Joudioux 2011] J. Joudioux, "Integral formula for the characteristic Cauchy problem on a curved background", *J. Math. Pures Appl.* (9) **95**:2 (2011), 151–193. MR 2012d:58044 Zbl 1223.35110

[Junker 1995] W. Junker, *Adiabatic vacua and Hadamard states for scalar quantum fields on curved spacetime*, Ph.D. thesis, Universität Hamburg, 1995. arXiv hep-th/9507097

[Junker and Schrohe 2002] W. Junker and E. Schrohe, "Adiabatic vacuum states on general spacetime manifolds: definition, construction, and physical properties", *Ann. Henri Poincaré* **3**:6 (2002), 1113–1181. MR 2004i:81172 Zbl 1038.81052

[Khavkine and Moretti 2015] I. Khavkine and V. Moretti, *Algebraic QFT in curved spacetime and quasifree Hadamard states: an introduction*, edited by R. Brunetti et al., Springer, Cham, 2015.

[Moretti 2006] V. Moretti, "Uniqueness theorem for BMS-invariant states of scalar QFT on the null boundary of asymptotically flat spacetimes and bulk-boundary observable algebra correspondence", *Comm. Math. Phys.* **268**:3 (2006), 727–756. MR 2008b:81193 Zbl 1141.81028

[Moretti 2008] V. Moretti, "Quantum out-states holographically induced by asymptotic flatness: invariance under spacetime symmetries, energy positivity and Hadamard property", *Comm. Math. Phys.* **279**:1 (2008), 31–75. MR 2009b:83055 Zbl 1145.83016

[Nicolas 2002] J.-P. Nicolas, *Dirac fields on asymptotically flat space-times*, Dissertationes Math. **408**, Instytut Matematyczny Polskiej Akademii Nauk, Warsaw, 2002. MR 2004c:83068 Zbl 1011.83015

[Nicolas 2006] J.-P. Nicolas, "On Lars Hörmander's remark on the characteristic Cauchy problem", *Ann. Inst. Fourier* (*Grenoble*) **56**:3 (2006), 517–543. MR 2008d:35114 Zbl 1124.35037

[Radzikowski 1996] M. J. Radzikowski, "Micro-local approach to the Hadamard condition in quantum field theory on curved space-time", *Comm. Math. Phys.* **179**:3 (1996), 529–553. MR 97f:81107 Zbl 0858.53055

[Rodino 1975] L. Rodino, "A class of pseudo differential operators on the product of two manifolds and applications", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **2**:2 (1975), 287–302. MR 53 #1656 Zbl 0308.35081

[Ruzhansky and Turunen 2010] M. Ruzhansky and V. Turunen, *Pseudo-differential operators and symmetries: background analysis and advanced topics*, Pseudo-Differential Operators: Theory and Applications **2**, Birkhäuser, Basel, 2010. MR 2011b:35003 Zbl 1193.35261

[Sahlmann and Verch 2001] H. Sahlmann and R. Verch, "Microlocal spectrum condition and Hadamard form for vector-valued quantum fields in curved spacetime", *Rev. Math. Phys.* **13**:10 (2001), 1203–1246. MR 2002m:81169 Zbl 1029.81053

[Sanders 2010] K. Sanders, "Equivalence of the (generalised) Hadamard and microlocal spectrum condition for (generalised) free fields in curved spacetime", *Comm. Math. Phys.* **295**:2 (2010), 485–501. MR 2011g:81167 Zbl 1192.53072

[Wald 1984] R. M. Wald, *General relativity*, University of Chicago Press, 1984. MR 86a:83001 Zbl 0549.53001

[Wrochna 2013] M. Wrochna, *Singularities of two-point functions in quantum field theory*, Ph.D. thesis, Georg-August-Universität, Göttingen, 2013, available at http://hdl.handle.net/11858/00-1735-0000-0001-BB3C-E.

Received 8 Nov 2014. Revised 13 Oct 2015. Accepted 16 Nov 2015.

CHRISTIAN GÉRARD: christian.gerard@math.u-psud.fr
Laboratoire de Mathématiques d'Orsay, Université Paris-Sud XI, 91405 Orsay, France

MICHAŁ WROCHNA: michal.wrochna@ujf-grenoble.fr
Institut Joseph Fourier, Université Joseph Fourier (Grenoble 1) UMR 5582 CNRS, 38402 Grenoble, France

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.151

GLOBAL-IN-TIME STRICHARTZ ESTIMATES ON NONTRAPPING, ASYMPTOTICALLY CONIC MANIFOLDS

ANDREW HASSELL AND JUNYONG ZHANG

We prove global-in-time Strichartz estimates without loss of derivatives for the solution of the Schrödinger equation on a class of nontrapping asymptotically conic manifolds. We obtain estimates for the full set of admissible indices, including the endpoint, in both the homogeneous and inhomogeneous cases. This result improves on the results by Tao, Wunsch and the first author and by Mizutani, which are local in time, as well as results of the second author, which are global in time but with a loss of angular derivatives. In addition, the endpoint inhomogeneous estimate is a strengthened version of the uniform Sobolev estimate recently proved by Guillarmou and the first author. The second author has proved similar results for the wave equation.

1.	Introduction	151
2.	Spectral measure and partition of the identity at low energies	157
3.	Spectral measure and partition of the identity at high energies	166
4.	Proof of Proposition 1.5	172
5.	L^2 estimates	176
6.	Dispersive estimates	180
7.	Homogeneous Strichartz estimates	183
8.	Inhomogeneous Strichartz estimates	184
	Acknowledgements	190
	References	190

1. Introduction

Strichartz estimates are an essential tool for studying the behaviour of solutions to nonlinear Schrödinger equations, nonlinear wave equations and other nonlinear dispersive equations. In particular, global-in-time Strichartz estimates are needed to show global well-posedness and scattering for these equations. The purpose of this article is to prove global-in-time Strichartz estimates for the Schrödinger equation on asymptotically conic, nontrapping manifolds.

Let (M°, g) be a Riemannian manifold of dimension $n \geq 2$ and let $I \subset \mathbb{R}$ be a time interval. Strichartz estimates are a family of dispersive estimates on solutions $u(t,z) : I \times M^\circ \to \mathbb{C}$ to the Schrödinger equation

$$i\,\partial_t u + \Delta_g u = 0, \quad u(0) = u_0(z), \tag{1-1}$$

MSC2010: primary 35Q41; secondary 58J40.
Keywords: Strichartz estimates, asymptotically conic manifolds, spectral measure, Schrödinger propagator.

where Δ_g denotes the Laplace–Beltrami operator on (M°, g). The general Strichartz estimates state that

$$\|u(t,z)\|_{L^q_t L^r_z(I\times M^\circ)} \le C\|u_0\|_{H^s(M^\circ)},$$

where H^s denotes the L^2-Sobolev space over M° and (q, r) is an *admissible pair*, i.e.,

$$2 \le q, r \le \infty, \quad \frac{2}{q} + \frac{n}{r} = \frac{n}{2}, \quad (q, r, n) \neq (2, \infty, 2). \tag{1-2}$$

It is well known that (1-1) holds for $(M^\circ, g) = (\mathbb{R}^n, \delta)$ with $s = 0$ and $I = \mathbb{R}$.

In this paper, we continue the investigations carried out in [Hassell et al. 2005; 2006] concerning Strichartz inequalities on a class of non-Euclidean spaces, that is, smooth, complete, noncompact, asymptotically conic Riemannian manifolds (M°, g) which satisfy a nontrapping condition. Here, "asymptotically conic" means that M° has an end of the form $(r_0, \infty)_r \times Y$, with metric asymptotic to $dr^2 + r^2h$ as $r \to \infty$, where (Y, h) is a closed Riemannian manifold of dimension $n-1$ (a more precise definition is given below). Hassell, Tao and Wunsch [Hassell et al. 2006] established the local-in-time Strichartz inequalities

$$\|e^{it\Delta_g}u_0\|_{L^q_t L^r_z([0,1]\times M^\circ)} \le C\|u_0\|_{L^2(M^\circ)}. \tag{1-3}$$

Here, we establish the same inequality on the full time interval $\mathbb{R}$. To treat an infinite time interval, the method of [Hassell et al. 2006] no longer works and we take a completely new approach (see Section 1C). Although phrased in terms of asymptotically conic manifolds, we emphasize that our results apply in particular to

- Schrödinger operators $\Delta + V$ on $\mathbb{R}^n$ with V suitably regular and decaying at infinity;
- nontrapping metric perturbations of flat Euclidean space with the perturbation suitably regular and decaying at infinity.

1A. ***Geometric setting.*** Let us recall the asymptotically conic geometric setting, which is the same as in [Guillarmou et al. 2013a; 2013b; Hassell and Wunsch 2005; Hassell et al. 2006]. Let (M°, g) be a complete, noncompact Riemannian manifold of dimension $n \ge 2$ with one end, diffeomorphic to $(0, \infty) \times Y$, where Y is a smooth, compact, connected manifold without boundary. Moreover, we assume (M°, g) is asymptotically conic, which means that M° can be compactified to a manifold M with boundary $\partial M = Y$ such that the metric g becomes a scattering metric on M. That is, in a collar neighbourhood $[0, \epsilon)_x \times \partial M$ of ∂M, g takes the form

$$g = \frac{dx^2}{x^4} + \frac{h(x)}{x^2} = \frac{dx^2}{x^4} + \frac{\sum h_{jk}(x, y)\, dy^j\, dy^k}{x^2}, \tag{1-4}$$

where $x \in C^\infty(M)$ is a boundary defining function for ∂M and h is a smooth family of metrics on Y. Here we use $y = (y_1, \dots, y_{n-1})$ for local coordinates on $Y = \partial M$ and the local coordinates (x, y) on M near ∂M. Away from ∂M, we use $z = (z_1, \dots, z_n)$ to denote the local coordinates. Moreover, if every geodesic $z(s)$ in M reaches Y as $s \to \pm\infty$, we say M is nontrapping. The function $r := 1/x$ near $x = 0$ can be thought of as a "radial" variable near infinity and $y = (y_1, \dots, y_{n-1})$ can be regarded as $n - 1$

"angular" variables. Rewriting (1-4) using coordinates (r, y), we see that the metric is asymptotic to the exact conic metric $dr^2 + r^2 h(0)$ on $(r_0, \infty)_r \times Y$ as $r \to \infty$.

The Euclidean space $M^\circ = \mathbb{R}^n$, or any compactly supported perturbation of this metric, is an example of an asymptotically conic manifold with Y equal to $\mathbb{S}^{n-1}$ endowed with the standard metric.

Let (M°, g) be an asymptotically conic manifold. The complex Hilbert space $L^2(M^\circ)$ is given by the inner product

$$\langle f_1, f_2 \rangle_{L^2(M^\circ)} = \int_{M^\circ} f_1(z)\overline{f_2(z)}\, dg(z),$$

where $dg(z) = \sqrt{g}\, dz$ is the measure induced by the metric g. Let $\Delta_g = \nabla^*\nabla$ be the Laplace–Beltrami operator on M; our sign convention is that Δ_g is a positive operator. Let V be a real potential function on M such that

$$V \in C^\infty(M), \qquad V(x, y) = O(x^3) \quad \text{as } x \to 0. \tag{1-5}$$

We assume that $n \geq 3$ and that one of two conditions hold: either

$$\boldsymbol{H} := \Delta_g + V \quad \text{has no zero eigenvalue or zero-resonance,} \tag{1-6}$$

or the stronger condition

$$\boldsymbol{H} := \Delta_g + V \quad \text{has no nonpositive eigenvalues or zero-resonance.} \tag{1-7}$$

By a zero-resonance we mean a nontrivial solution u to $\boldsymbol{H}u = 0$ such that $u \to 0$ at infinity. Notice that the second assumption, (1-7), implies that $\boldsymbol{H}$ is a nonnegative operator, so that we can define $\sqrt{\boldsymbol{H}}$. These assumptions allow us to use the results of [Guillarmou et al. 2013a; 2013b].

1B. ***Main results.*** Now we consider the Schrödinger equation

$$i\,\partial_t u + \boldsymbol{H}u = 0, \quad u(0, \cdot) = u_0 \in L^2(M). \tag{1-8}$$

The main purpose of this paper is to prove the following results. Notice that the endpoint estimate ($q = 2$ and $\tilde{q} = 2$) is included in both cases.

Theorem 1.1 (long-time homogeneous Strichartz estimate). *Let (M°, g) be an asymptotically conic, nontrapping manifold of dimension $n \geq 3$. Let $\boldsymbol{H} = \Delta_g + V$ satisfy* (1-5) *and* (1-7) *and suppose u is the solution to* (1-8). *Then*

$$\|u(t, z)\|_{L^q_t L^r_z(\mathbb{R}\times M^\circ)} \leq C\|u_0\|_{L^2(M^\circ)} \tag{1-9}$$

provided the admissible pair $(q, r) \in [2, \infty]^2$ satisfies (1-2).

Theorem 1.2 (long-time inhomogeneous Strichartz estimate). *Let (M°, g) and $\boldsymbol{H}$ be as in Theorem 1.1. Suppose that u solves the inhomogeneous Schrödinger equation with zero initial data*

$$i\,\partial_t u + \boldsymbol{H}u = F(t, z), \quad u(0, \cdot) = 0. \tag{1-10}$$

Then the inhomogeneous Strichartz estimate

$$\|u(t, z)\|_{L^q_t L^r_z(\mathbb{R}\times M^\circ)} \leq C\|F\|_{L^{\tilde{q}'}_t L^{\tilde{r}'}_z(\mathbb{R}\times M^\circ)} \tag{1-11}$$

holds for admissible pairs (q, r), $(\tilde{q}, \tilde{r})$.

Remark 1.3. If we make the weaker assumption (1-6), then the statements above still hold, provided that u_0 and $F(t, \cdot)$ lie in the positive spectral subspace of $\boldsymbol{H}$, or in other words that $u_0 = 1_{[0,\infty)}(\boldsymbol{H})(u_0)$, and similarly for $F(t, \cdot)$ for almost every t.

Remark 1.4. We restrict to $n \geq 3$ since the results of [Guillarmou and Hassell 2008] only apply to that case. More recently, Sher [2013] has extended these results to $n = 2$; using his results, one could treat the case $n = 2$ also (noting that the endpoint estimates fail in dimension 2, due to a logarithmic divergence in the resolvent at zero energy occurring in dimension 2). For space reasons, we have not attempted to treat this case in the present paper.

1C. ***Strategy of the proof.*** Our argument here extends to long time and to the endpoint Strichartz estimates of Hassell et al. [2006], who constructed a "local" parametrix for the propagator $e^{it\boldsymbol{H}}$ based on the parametrix from [Hassell and Wunsch 2005]. In that paper, Schrödinger solutions $e^{it\boldsymbol{H}}u_0$ were obtained by applying the parametrix to u_0 and then correcting this approximate solution using Duhamel's formula, using local smoothing estimates to control the correction term. This approach works well on a finite time interval, but cannot be expected to work on an infinite time interval as the errors accumulate over time; certainly they cannot be expected to decay to zero as $t \to \infty$, as would be required to prove L^q estimates in time on an infinite interval.

The main new idea in the current paper is to express the propagator $e^{it\boldsymbol{H}}$ *exactly*, using the spectral measure $dE_{\sqrt{\boldsymbol{H}}}(\lambda)$, exploiting the very precise information on the spectral measure for the Laplacian on asymptotically conic, nontrapping manifolds that has recently become available from the works [Hassell and Vasy 1999; Hassell and Wunsch 2008; Guillarmou et al. 2013a].

After expressing the propagator in terms of an integral of the multiplier $e^{it\lambda^2}$ against the spectral measure, our strategy is to use the abstract Strichartz estimate proved in [Keel and Tao 1998]. Thus, with $U(t)$ denoting the (abstract) propagator, we need to show uniform $L^2 \to L^2$ estimates for $U(t)$, and a $L^1 \to L^\infty$ type dispersive estimate on the $U(t)U(s)^*$ with a bound of the form $O(|t-s|^{-n/2})$. In the flat Euclidean setting, the estimates are obvious because of the explicit formula for the propagator. But in our general setting it turns out to be more complicated. It follows from [Hassell and Wunsch 2005] that the propagator $U(t)(z, z')$ fails to satisfy such a dispersive estimate at any pair of conjugate points $(z, z') \in M^\circ \times M^\circ$ (i.e., pairs (z, z') where geodesics emanating from z focus at z'). Our geometric assumptions allow conjugate points, so we need to modify the propagator such that the failure of the dispersive estimate at conjugate points is avoided.

This is possible due to the TT^* nature of the estimates required by the Keel–Tao formalism. Recall that the dispersive estimate required by Keel and Tao is of the form

$$\|U(t)U(s)^*\|_{L^1 \to L^\infty} \leq C|t-s|^{-n/2}. \tag{1-12}$$

If $U(t)$ is the propagator $e^{it\boldsymbol{H}}$ then the operator on the left-hand side is $e^{i(t-s)\boldsymbol{H}}$. However, nothing in the Keel–Tao formalism requires the $U(t)$ to form a group of operators. Hence we are free to break up $e^{it\boldsymbol{H}} = \sum_j U_j(t)$ and prove the estimate (1-12) for each U_j. Our choice of $U_j(t)$ (sketched directly below) means that $U_j(t)U_j(s)^*$ is essentially the kernel $e^{i(t-s)\boldsymbol{H}}$ localized sufficiently close to the diagonal that we avoid pairs of conjugate points, and hence can prove the dispersive estimate.

Our method of decomposing $e^{it\boldsymbol{H}} = \sum_j U_j(t)$ is motivated by a decomposition used in the proof in [Guillarmou et al. 2013b] of a *restriction estimate* for the spectral measure, that is, an estimate of the form

$$\|dE_{\sqrt{\boldsymbol{H}}}(\lambda)\|_{L^p(M^\circ)\to L^{p'}(M^\circ)} \le C\lambda^{n(1/p-1/p')-1}, \quad 1 \le p \le \frac{2(n+1)}{n+3}.$$

In [Guillarmou et al. 2013b], it was observed that, to prove a restriction estimate for $dE_{\sqrt{\boldsymbol{H}}}(\lambda)$, it suffices (via a TT^* argument) to prove the same estimate for the operators $Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j(\lambda)^*$, where $Q_j(\lambda)$ is a partition of the identity operator in $L^2(M^\circ)$. The operators $Q_j(\lambda)$ used in [Guillarmou et al. 2013b] are pseudodifferential operators (of a certain specific type) serving to localize $dE_{\sqrt{\boldsymbol{H}}}(\lambda)$ in phase space close to the diagonal. Guillarmou et al. [2013b] showed that the localized operators $Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j(\lambda)^*$ satisfy kernel estimates analogous to those satisfied by the spectral measure for $\sqrt{\Delta}$ on flat Euclidean space:

$$\big|(Q_j(\lambda)\, dE^{(l)}_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j(\lambda))(z,z')\big| \le C\lambda^{n-1-l}(1+\lambda d(z,z'))^{-(n-1)/2+l}, \quad l \in \mathbb{N}, \tag{1-13}$$

where $dE^{(l)}_{\sqrt{\boldsymbol{H}}}(\lambda)$ is the l-th derivative in λ of the spectral measure and d is the Riemannian distance on M°.

The authors of [Guillarmou et al. 2013b] hoped that (1-13) could be used as a "black box" in applications of their work. Unfortunately, (1-13) seems inadequate for our present purposes. This is because, in order to obtain the dispersive estimate, we need to efficiently exploit the oscillation of the "spectral multiplier" $e^{it\lambda^2}$, and particularly the discrepancy between the way this function oscillates relative to the oscillations (in λ) of the Schwartz kernel of the spectral measure. The second main innovation of this paper is to improve the estimate (1-13) on the localized spectral measure. We show:

Proposition 1.5. *Let (M°, g) and $\boldsymbol{H}$ be as in Theorem 1.1. Then there exists a λ-dependent operator partition of unity on $L^2(M)$*

$$\mathrm{Id} = \sum_{j=1}^{N} Q_j(\lambda),$$

with N independent of λ, such that for each $1 \le j \le N$ we can write

$$(Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j^*(\lambda))(z,z') = \lambda^{n-1}\Big(\sum_{\pm} e^{\pm i\lambda d(z,z')} a_\pm(\lambda,z,z') + b(\lambda,z,z')\Big), \tag{1-14}$$

with estimates

$$|\partial_\lambda^\alpha a_\pm(\lambda,z,z')| \le C_\alpha \lambda^{-\alpha}(1+\lambda d(z,z'))^{-(n-1)/2}, \tag{1-15}$$

$$|\partial_\lambda^\alpha b(\lambda,z,z')| \le C_{\alpha,M} \lambda^{-\alpha}(1+\lambda d(z,z'))^{-K} \quad \textit{for any } K. \tag{1-16}$$

Here, $d(\cdot,\cdot)$ is the Riemannian distance on M°.

Remark 1.6. The estimates (1-15)–(1-16) are easily seen to imply (1-13) (using Lemma 2.3 to estimate the λ-derivatives of the operators $Q_i(\lambda)$). However, (1-15)–(1-16) also capture the oscillatory behaviour of the spectral measure, which is crucial in obtaining sharp dispersive estimates in Section 6.

We now define localized (in phase space) propagators $U_j(t)$ by

$$U_j(t) = \int_0^\infty e^{it\lambda^2} Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda), \quad 1 \le j \le N. \tag{1-17}$$

Then the operator $U_j(t)U_j(s)^*$ is given, at least formally, by (see Lemma 5.3)

$$U_j(t)U_j(s)^* = \int e^{i(t-s)\lambda^2} Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_j(\lambda)^*. \tag{1-18}$$

However, there are subtleties involved in spectral integrals such as (1-17)–(1-18) containing operator-valued functions. Even to show that (1-17) is well-defined as a bounded operator on $L^2(M^\circ)$ is nontrivial. The third main innovation of this paper is to give an effective method for analyzing spectral integrals such as (1-17)–(1-18) with operator-valued multipliers. We use a dyadic decomposition in λ and a Cotlar–Stein almost orthogonality argument to show the well-definedness of (1-17) and prove a uniform estimate on $\|U_j(t)\|_{L^2\to L^2}$, as required by the Keel–Tao formalism.

Having made sense of (1-18), we exploit the oscillations both in the multiplier $e^{i(t-s)\lambda^2}$ and in the localized spectral measure (as expressed by (1-15)–(1-16)) to obtain the required dispersive estimate for $U_j(t)U_j(s)^*$. The homogeneous Strichartz estimate for e^{itH} then follows by applying Keel–Tao to each U_j and summing over j.

Next we consider the inhomogeneous Strichartz estimates. As is well known, the non-endpoint cases of the inhomogeneous estimate follow from the homogeneous estimate and the Christ–Kiselev lemma. The endpoint inhomogeneous estimate requires an additional argument and, in particular, in this case we require estimates on $U_i(t)U_j(s)^*$ for $i \ne j$. This estimate turns out to be very similar to the uniform Sobolev estimate (on asymptotically conic, nontrapping manifolds) of Guillarmou and Hassell [2014]. We use the techniques of that paper, in particular a refined partition of the identity operator. This resemblance to their proof is more than formal: as pointed out to us by Thomas Duyckaerts and Colin Guillarmou, the inhomogeneous endpoint Strichartz estimate implies the uniform Sobolev estimate; we sketch this argument in Section 8. Thus, this part of the paper can be regarded as a time-dependent reformulation of the proof in [Guillarmou and Hassell 2014], leading to a more general result.

1D. ***Previous literature.*** Now we review some classical results about the Strichartz estimates. In the flat Euclidean space, where $M^\circ = \mathbb{R}^n$ and $g_{jk} = \delta_{jk}$, one can take $I = \mathbb{R}$; see [Strichartz 1977; Ginibre and Velo 1985; Keel and Tao 1998] and references therein. The now-classic paper [Keel and Tao 1998] developed an abstract approach to Strichartz estimates, which has become the standard approach in most subsequent literature, including this paper. Strichartz estimates for compact metric perturbations of Euclidean space were proved locally in time by Staffilani and Tataru [2002] and subsequently for asymptotically Euclidean manifolds by Robbiano and Zuily [2005] and Bouclet and Tzvetkov [2007], and in the asymptotically conic setting by Hassell et al. [2006] and Mizutani [2012]. In these works, either the metric is assumed to be nontrapping, or the theorem holds outside a compact set. Burq et al. [2010] proved that Strichartz estimates without loss hold on an asymptotically conic manifold with hyperbolic trapped set. Strichartz estimates have also been studied on exact cones [Ford 2010] and on asymptotically hyperbolic spaces [Bouclet 2011].

There has also been work on Strichartz estimates on compact manifolds and on manifolds with boundary. In the compact case, Strichartz estimates usually are local in time and with some loss of derivatives s (i.e., the right-hand side of (1-9) has to be replaced by the H^s norm of u_0). Estimates for the standard flat 2-torus were shown by Bourgain [1999] to hold for any $s > 0$. For any compact manifold, Burq et al. [2004a] showed that the estimate holds for $s = 1/q$ and that the loss of derivatives, as well as the localization in time, is sharp on the sphere. Manifolds with boundary were studied in [Blair et al. 2008; 2009; 2012; Ivanovici 2010].

Global-in-time Strichartz estimates on asymptotically Euclidean spaces have been proved by Bouclet and Tzvetkov [2008] (but with a low energy cutoff), Metcalfe and Tataru [2012], Marzuola, Metcalfe and Tataru [Marzuola et al. 2008] and Marzuola, Metcalfe, Tataru and Tohaneanu [Marzuola et al. 2010].

The second author has obtained global-in-time Strichartz estimates for the wave equation on asymptotically conic nontrapping manifolds [Zhang 2015b] and for the Schrödinger equation [Zhang 2015a].

As already noted, Strichartz estimates are an essential tool for studying the behaviour of solutions to nonlinear dispersive equations. There is a vast literature on this topic, and it is beyond the scope of this introduction to review it, so we refer instead to Tao's book [2006] and the references therein.

1E. ***Organization of this paper.*** We review the partition of the identity and properties of the microlocalized spectral measure for low energies in Section 2 and for high frequency in Section 3. In Section 4, we prove Proposition 1.5 based on the properties of the microlocalized spectral measure. Section 5 is devoted to the construction of microlocalized propagators and the proof of the L^2 estimates. The dispersive estimates are proved in Section 6. Finally, we prove the homogeneous Strichartz estimates in Section 7 and the inhomogeneous Strichartz estimates in Section 8.

2. Spectral measure and partition of the identity at low energies

The spectral measure for the operator $\boldsymbol{H}$ for low energies was constructed in [Guillarmou and Hassell 2008] on the "low energy space" $M^2_{k,b}$. Here we recall the low energy space $M^2_{k,b}$ and the associated space $M^2_{k,\mathrm{sc}}$. The latter space is needed in order to define the class of pseudodifferential operators in which our operator partition $Q_j(\lambda)$ from Proposition 1.5 lies.

2A. ***Low energy space.*** The low energy space $M^2_{k,b}$, defined in [Guillarmou and Hassell 2008] (based on unpublished work of Melrose and Sá Barreto), is a blown-up version of[1] $[0, \lambda_0] \times M^2$. This space is illustrated in Figure 1. More precisely, we define the codimension-3 corner $C_3 = \{0\} \times \partial M \times \partial M$ and the codimension-2 submanifolds

$$C_{2,L} = \{0\} \times \partial M \times M, \quad C_{2,R} = \{0\} \times M \times \partial M, \quad C_{2,C} = [0, 1] \times \partial M \times \partial M.$$

[1]In [Guillarmou and Hassell 2008], the spectral parameter was denoted by k rather than λ, hence the subscript "k" in $M^2_{k,b}$.

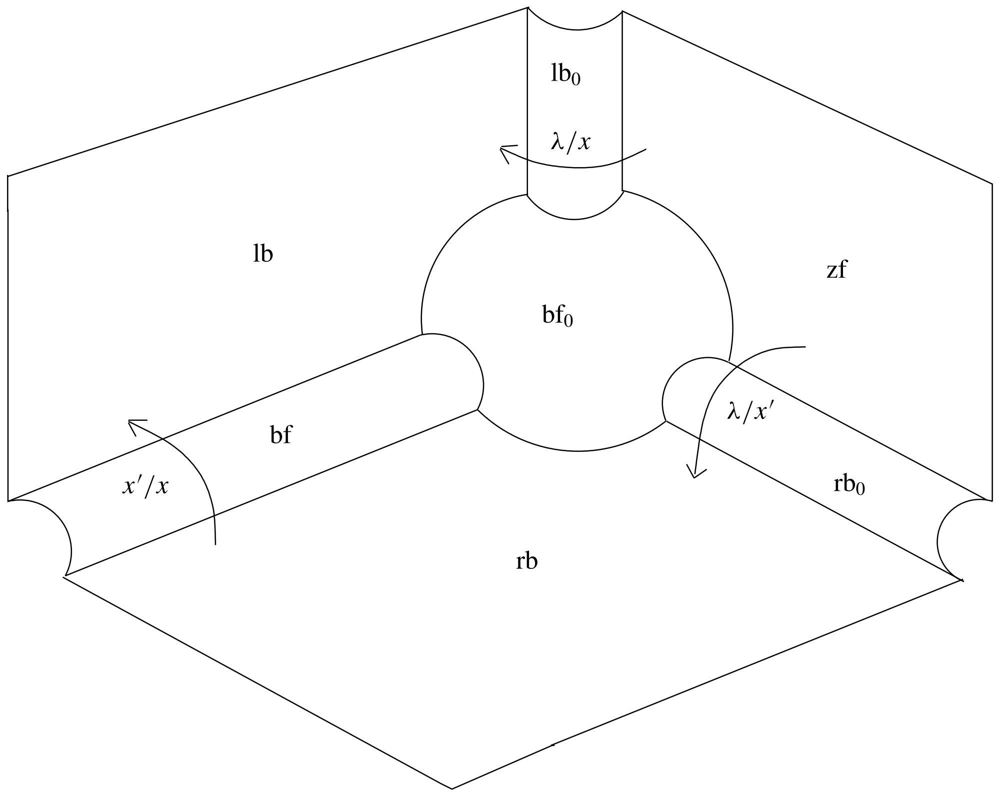

Figure 1. The manifold $M^2_{k,b}$. Arrows show the direction in which the indicated function increases from 0 to ∞.

Without loss of generality, we assume $\lambda_0 = 1$. The space $M^2_{k,b}$ is defined by

$$M^2_{k,b} = [[0, 1] \times M^2; C_3, C_{2,R}, C_{2,L}, C_{2,C}]$$

with blow-down map $\beta_{\mathrm{b}} : M^2_{k,b} \to [0, 1] \times M^2$. Here the notation $[X; Y]$, where X is a manifold with corners and Y a p-submanifold of X,[2] indicates that Y is blown up in X in the real sense; as a set, $[X; Y]$ is the disjoint union of $X \setminus Y$ and the inward-pointing spherical normal bundle SN^+Y of Y. Moreover, $[X; Y_1, Y_2, \dots]$ indicates iterated blow-up. See [Melrose 1994, Section 18] for further details.

The new boundary hypersurfaces created by these blow-ups are labelled by

$$\mathrm{rb} = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}([0, 1] \times M \times \partial M), \quad \mathrm{lb} = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}([0, 1] \times \partial M \times M), \quad \mathrm{zf} = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}(\{0\} \times M \times M),$$

the “b-face” $\mathrm{bf} = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}(C_{2,C} \setminus C_3)$ and

$$\mathrm{bf}_0 = \beta_{\mathrm{b}}^{-1}(C_3), \quad \mathrm{rb}_0 = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}(C_{2,R} \setminus C_3), \quad \mathrm{lb}_0 = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}(C_{2,L} \setminus C_3).$$

[2]We say that Y is a p-submanifold of X if, near every point $p \in Y$, there are local coordinates $x_1, \dots, x_l, y_1, \dots, y_{n-l}$, where $x_i \geq 0$, $y_i \in (-\epsilon, \epsilon)$ and $p = (0, \dots, 0)$, such that Y is given locally by the vanishing of some subset of these coordinates.

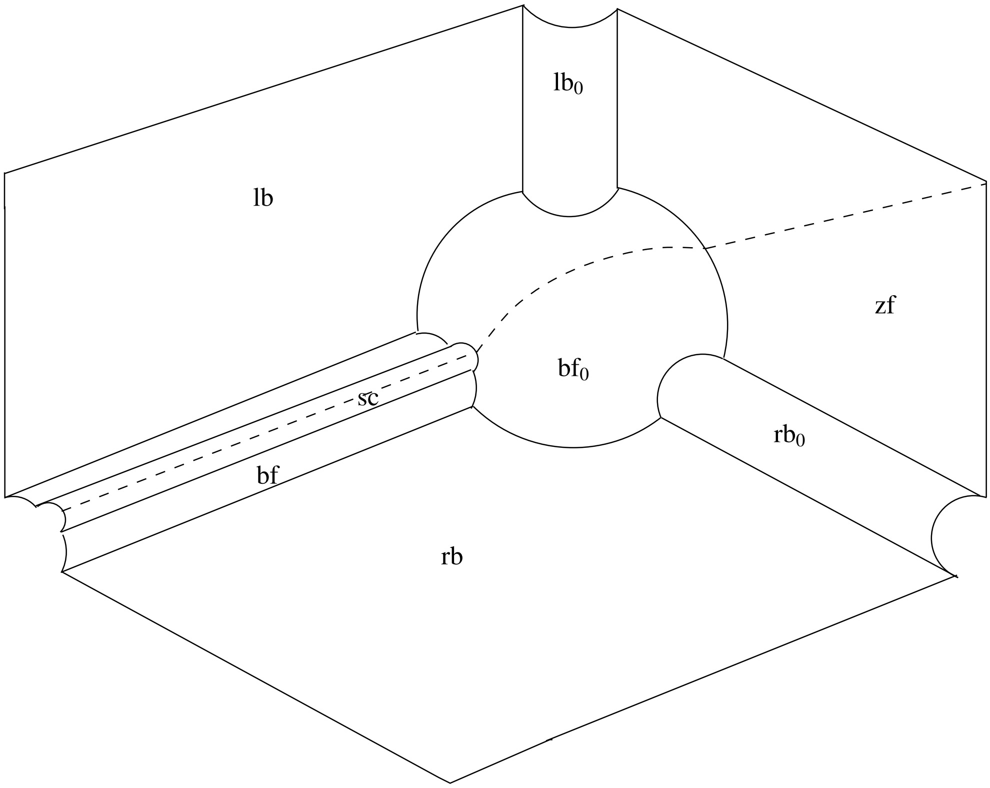

Figure 2. The manifold $M^2_{k,\mathrm{sc}}$; the dashed line is the boundary of the lifted diagonal $\Delta_{k,\mathrm{sc}}$.

The closed lifted diagonal is given by $\mathrm{diag_b} = \mathrm{clos}\, \beta_{\mathrm{b}}^{-1}([0,1] \times \{(m,m); m \in M^\circ\})$ and its intersection with the face bf is denoted by $\partial_{\mathrm{bf}}\, \mathrm{diag_b}$. We remark that zf is canonically diffeomorphic to the b-double space

$$M^2_{\mathrm{b}} = [M^2; \partial M \times \partial M], \tag{2-1}$$

as is each section $M^2_{k,b} \cap \{\lambda = \lambda_*\}$ for fixed $0 < \lambda_* < 1$.

We further define the space $M^2_{k,\mathrm{sc}}$ to be the blow-up of $M^2_{k,b}$ at $\partial_{\mathrm{bf}}\, \mathrm{diag}_b$. This space is illustrated in Figure 2. The sections $M^2_{k,\mathrm{sc}} \cap \{\lambda = \lambda_*\}$ for fixed $0 < \lambda_* < 1$ are all canonically diffeomorphic to the scattering double space M^2_{sc}, which is the blow-up of M^2_{b} at the boundary of the lifted diagonal

$$M^2_{\mathrm{sc}} = [M^2_{\mathrm{b}}; \partial\, \mathrm{diag}_b].$$

To avoid excessive notation, we denote the diagonal in M^2_{b} and in $M^2_{k,b}$ by the same symbol diag_b. We similarly define $\mathrm{diag_{sc}}$ to be the closure of the interior of diag_b lifted to M^2_{sc} (or $M^2_{k,\mathrm{sc}}$).

2B. ***Coordinates.*** Let $(x, y) = (x, y_1, \dots, y_{n-1})$ be local coordinates on M near a boundary point, as discussed in Section 1A. We define functions x and y on $M^2_{k,b}$ by lifting from the left factor of M (near ∂M), and x' and y' by lifting from the right factor of M; and similarly z and z' (away from ∂M).

Let $\rho = x/\lambda$, $\rho' = x'/\lambda$ and $\sigma = \rho/\rho' = x/x'$. Then we can use coordinates $(y, y', \sigma, \rho', \lambda)$ near bf and away from rb, while $(y, y', \sigma^{-1}, \rho, \lambda)$ near bf and away from lb.

Next we consider local coordinates on the scattering double space M^2_{sc}. The only difference between this space and M^2_{b} is at the boundary of the diagonal. In local coordinates, near $\partial_{\mathrm{bf}}\,\mathrm{diag}_b$, a boundary defining function for bf is given by x/λ and the diagonal is given by $\sigma = 1,\ y = y'$. Therefore, coordinates on the interior of the new boundary hypersurface, denoted by sc, created by this blow-up are

$$\frac{\lambda(\sigma - 1)}{x}, \quad \frac{\lambda(y - y')}{x}, \quad \lambda, \quad y'.$$

We also need to consider coordinates on phase space. As emphasized by Melrose [1994], the appropriate phase space for analyzing the Laplacian with respect to a scattering metric is the scattering cotangent bundle. This is the dual space of the scattering tangent bundle ${}^{\mathrm{sc}}TM$, which is the bundle whose sections are the smooth vector fields over M which are uniformly of finite length with respect to g. Near the boundary, due to the form of the metric (1-4), they are spanned over $C^\infty(M)$ by the vector fields $x^2\,\partial_x$ and $x\,\partial_{y_i}$. Dually, the scattering cotangent bundle is spanned near the boundary by vector fields $dx/x^2 = -d(1/x)$ and dy_i/x; away from the boundary, it is canonically diffeomorphic to the usual cotangent bundle. Thus, a point in the scattering cotangent bundle can be expressed as a linear combination

$$\nu\lambda d\left(\frac{1}{x}\right) + \sum_{i=1}^{n-1} \lambda\mu_i\,\frac{dy_i}{x} \tag{2-2}$$

near the boundary, or

$$\sum_{i=1}^{n} \lambda\zeta_i\,dz_i \tag{2-3}$$

away from the boundary, which defines linear coordinates (μ, ν) or ζ on each fibre of the scattering cotangent bundle. Notice that we have introduced a scaling by the spectral parameter λ; as $\lambda = 1/h$, this is essentially the semiclassical scaling, appropriate to our operator $\Delta - \lambda^2 = \lambda^2(h^2\Delta - 1)$, although in this low energy case, we are looking at the limit $h \to \infty$, rather than $h \to 0$ as in the high energy case in Section 3.

The appropriate "compressed cotangent bundle" over $M^2_{k,b}$ is discussed in [Guillarmou et al. 2013a, Section 2.3]. Here, we only describe this for $\lambda > 0$ plus a neighbourhood of the boundary hypersurface bf. In this region, it is given by the lift of the bundle ${}^{\mathrm{sc}}T^*M \times {}^{\mathrm{sc}}T^*M$ to $M^2 \times [0, 1]$ and then to $M^2_{k,b}$. In particular, we use coordinates (μ, ν) lifted from the left factor of M and (μ', ν') lifted from the right factor of M in a neighbourhood of bf. We remark that these coordinates remain valid in a neighbourhood of bf even at $\lambda = 0$, which follows from the fact that (2-2) can be written in the form

$$\nu d\left(\frac{1}{\rho}\right) + \sum_{i=1}^{n-1} \mu_i\,\frac{dy_i}{\rho}.$$

The following lemma will be useful in our estimates in Section 4.

Lemma 2.1. *Let $w = (w_1, \dots, w_n)$ denote a set of defining functions for* $\mathrm{diag}_b \subset M^2_{k,b}$; *that is, the differentials dw_i are linearly independent and* $\mathrm{diag}_b = \{w = 0\}$. *For example, near* bf_0 *or* bf, *we can take $w = (\sigma - 1, y_1 - y_1', \dots, y_{n-1} - y_{n-1}')$. Then $|w|/x$ is comparable to $d(z, z')$ in a neighbourhood of* diag_b. *Equivalently, $|w|/\rho$ is comparable to $\lambda d(z, z')$.*

Proof. Away from $\mathrm{bf}_0 \cup \mathrm{bf}$, $|w|^2$ is a quadratic defining function for diag_b and so is $d(z, z')^2$, hence they are comparable. Now consider what happens near bf_0 or bf. In coordinates $w = (\sigma - 1, y_1 - y_1', \dots y_{n-1} - y_{n-1}')$, we have

$$\frac{|w|}{x} \sim \left|\frac{\sigma - 1}{x}\right| + \left|\frac{y - y'}{x}\right|.$$

Write $r = 1/x$; then this is

$$|r - r'| + r|y - y'|.$$

Given that the metric takes the form $dr^2 + r^2 h(x, y, dy)$, where h is positive definite, we see that this is comparable to $d(z, z')$. □

Remark 2.2. In the case $M^\circ = \mathbb{R}^n$, with Euclidean coordinates $z = (z_1, \dots, z_n)$, we can take $w = (z_1 - z_1', \dots, z_n - z_n')$.

2C. ***Pseudodifferential operators on the low energy space.*** We use the class of pseudodifferential operators $\Psi_k^m(M; \Omega_{k,b}^{1/2})$ on $M^2_{k,\mathrm{sc}}$ introduced by Guillarmou and Hassell [2008]. By definition, these operators have Schwartz kernels which are half-densities conormal to the diagonal $\mathrm{diag}_{\mathrm{sc}}$, smooth on $M^2_{k,\mathrm{sc}}$ away from the diagonal, and rapidly decreasing at all boundary hypersurfaces not meeting the diagonal, i.e., at lb_0, rb_0, lb and rb. In addition, we will only consider those operators with kernels supported where $\rho, \rho' \le C < \infty$. In this setting we can write the kernel in the form

$$\lambda^n \int e^{i\lambda/x\,((1-\sigma)\nu + (y - y')\cdot\mu)} a(\lambda, \rho, y, \mu, \nu)\, d\mu\, d\nu\, |dg\, dg'|^{1/2}, \tag{2-4}$$

where a is a classical symbol of order m in the (μ, ν) variables, smooth in (λ, ρ, y) and supported where $\rho \le c$. If we write this in the form $A(z, z', \lambda)\, |dg\, dg'|^{1/2}$, then the action on a half-density $f\, |dg|^{1/2}$ is given by

$$\left(\int A(z, z', \lambda) f(z')\, dg(z')\right) |dg(z)|^{1/2}.$$

Given that we have a canonical half-density factor, namely the Riemannian half-density $|dg|^{1/2}$, we will usually omit the half-density factors below.

From the representation (2-4) it is easy to see the following:

Lemma 2.3. *If $A \in \Psi_k^m(M; \Omega_{k,b}^{1/2})$ then $(\lambda\, \partial_\lambda)^N A$ is also a pseudodifferential operator of order m, i.e., $(\lambda \partial_\lambda)^N A \in \Psi_k^m(M; \Omega_{k,b}^{1/2})$.*

Proof. It suffices to prove this for $N = 1$ and use induction. If $\lambda\, \partial_\lambda$ hits the function a in (2-4), then a is still a symbol of order m in the (μ, ν) variables, smooth in (λ, ρ, y) and supported where $\rho \le c$. (Notice that $\rho = x/\lambda$ depends on λ as well.) On the other hand, if $\lambda\, \partial_\lambda$ hits the phase, this is the same as

$\nu\,\partial_\nu + \mu\,\partial_\mu$ hitting the phase, as it is homogeneous of degree 1 in both λ and (ν, μ). Integrating by parts, we obtain another symbol $\tilde{a}$ of order m. This completes the proof. □

Lemma 2.4. *If $A = A(z, z', \lambda)\,|dg\,dg'|^{1/2} \in \Psi_k^m(M; \Omega_{k,b}^{1/2})$ and $m < -n$, then A satisfies a kernel bound*

$$|A(z, z', \lambda)| \le \lambda^n (1 + \lambda d(z, z'))^{-N} \tag{2-5}$$

for any $N \in \mathbb{N}$.

Proof. If the order m is less than $-n$, then the integral (2-4) is absolutely convergent, showing that the kernel of $\lambda^{-n} A$ is uniformly bounded. Next, we note that the differential operator

$$\frac{1 - \partial_\nu^2 - \sum_i \partial_{\mu_i}^2}{1 + \lambda^2 (x^{-2}(\sigma - 1)^2 + x^{-2}|y - y'|^2)}$$

leaves the exponential in (2-4) invariant. By applying this N times to the exponential and then integrating by parts, we see that the integral is bounded by

$$C_N \bigl(1 + \lambda^2 (x^{-2}(\sigma - 1)^2 + x^{-2}|y - y'|^2)\bigr)^{-N}$$

for any N. Finally, as in the proof of Lemma 2.1, the square of the Riemannian distance on M is comparable to

$$\frac{(\sigma - 1)^2}{x^2} + \frac{|y - y'|^2}{x^2},$$

so the integral is bounded by $C_N (1 + \lambda\, d(z, z'))^{-N}$ for any N. □

Corollary 2.5. *If $A \in \Psi_k^m(M; \Omega_{k,b}^{1/2})$ and $m < -n$, then A is bounded $L^2(M^\circ) \to L^2(M^\circ)$ uniformly as $\lambda \to 0$. The same is true for $(\lambda\,\partial_\lambda)^N A$ for any N.*

Proof. This follows from the kernel bound in Lemma 2.4, the volume estimate $cr^n \le V(z, r) \le Cr^n$ for the volume $V(z, r)$ of the ball of radius r centred at $z \in M^\circ$, and Schur's test. □

2D. *Low energy partition of the identity.* Recall that, in Proposition 1.5, we employ a partition of the identity. We use essentially the same partition of the identity as in [Guillarmou et al. 2013b]. To define it, we specify the symbols of these operators, which must form a partition of unity on the phase space. We point out that, in our approach, it is crucial to be able to localize in phase space (and hence necessary to use pseudodifferential operators) in order to eliminate difficulties with conjugate points.

For low energies and for a given small positive ϵ, this partition is defined as follows. We first form an open cover $G_0 \cup \cdots \cup G_{N_l}$ of the phase space ${}^{sc}T^*M$. The set G_0 consists of all points away from the boundary, that is, the points with $x > \epsilon$. The next set G_1 consists of points near the boundary, say $x < 2\epsilon$, but away from the characteristic variety, that is, satisfying $|\mu|_h^2 + \nu^2 < \frac{1}{2}$ or $|\mu|_h^2 + \nu^2 > \frac{3}{2}$. We then break up the set $\bigl\{x < 2\epsilon,\ |\mu|_h^2 + \nu^2 \in \bigl[\frac{1}{4}, 2\bigr]\bigr\}$ into a finite number of sets $G_2, \dots, G_{N_l}$ such that, for each set G_j, the value of ν lies in some interval of length $\le \delta$, where δ is taken to be sufficiently small.

We then form a partition of unity subordinate to this open cover and take these as the principal symbols of pseudodifferential operators Q_j^{low} in the class $\Psi_k^0(M; \Omega_{k,b}^{1/2})$ described above. More precisely, we choose a function $\chi \in C^\infty(\mathbb{R})$ of a real variable with $\chi(t) = 0$ for $t \le \epsilon$ and $\chi(t) = 1$ for $t \ge 2\epsilon$. We define $Q_0^{\mathrm{low}}(\lambda)$ to be multiplication by the function $1-\chi(\rho)$ (recall $\rho = x/\lambda$). Next, we choose $Q_1'(\lambda)$ such that its (full) symbol is equal to 0 for $\frac{1}{2} \le |\mu|_h^2 + \nu^2 \le \frac{3}{2}$ and equal to 1 outside $\frac{1}{4} \le |\mu|_h^2 + \nu^2 \le 2$. Then we define $Q_1^{\mathrm{low}} = \chi(\rho) Q_1'$. This means that the symbol of $\mathrm{Id} - Q_0^{\mathrm{low}} - Q_1^{\mathrm{low}}$ is supported where ρ is small and close to the characteristic variety $|\mu|_h^2 + \nu^2 = 1$. We then decompose this as $Q_2^{\mathrm{low}} + \cdots + Q_{N_l}^{\mathrm{low}}$ so that the symbol of each Q_j^{low}, $j \ge 2$ in supported in G_j, hence supported where ν is contained in an interval of length $\le \delta$.

2E. ***Localized spectral measure.*** The main result of [Guillarmou et al. 2013a] was that the spectral measure for the Laplacian on an asymptotically conic manifold is, for low energies, a Legendre distribution associated to a pair of Legendre submanifolds, the "propagating Legendrian" L^{bf} and the "incoming/outgoing Legendrian" $L^{\sharp}$. We now explain very briefly what this means. We first have to introduce the contact manifold in which these Legendre submanifolds live. Consider the bundle ${}^{\Phi}T^*M_{\mathrm{b}}^2$, obtained by lifting ${}^{\mathrm{sc}}T^*M \times {}^{\mathrm{sc}}T^*M$ (viewed as a bundle over M^2) to M_{b}^2. This bundle carries a symplectic structure, but the symplectic form degenerates at the boundary. Nevertheless, it determines a contact structure on this bundle restricted to the boundary hypersurface bf,[3] which we denote by ${}^{\Phi}T^*_{\mathrm{bf}}M_{\mathrm{b}}^2$. We give this contact structure in local coordinates $(y, y', \sigma, \mu, \mu', \nu, \nu')$ for ${}^{\Phi}T^*_{\mathrm{bf}}M_{\mathrm{b}}^2$, where $\sigma = x/x'$, (μ, ν) are as in (2-2) and, as above, the unprimed/primed coordinates are lifted from the left/right copies of ${}^{\mathrm{sc}}T^*M$. In these coordinates, the contact form has an expression

$$d\nu - \mu \cdot dy + \sigma(d\nu' - \mu' \cdot dy').$$

A Legendrian submanifold is, by definition, an $2n-1$-dimensional submanifold of this $4n-1$-dimensional space on which the contact form vanishes. The Legendre submanifold $L^{\sharp}$ is easy to define: it is the submanifold

$$\{(y, y', \sigma, \mu, \mu', \nu, \nu') \mid \mu = \mu' = 0,\ \nu = \nu' = 1\}.$$

The other Legendre submanifold, L^{bf}, is more interesting. It encodes the geodesic flow on the cone over $(\partial M, h)$ where $h = h(0)$ is the metric in (1-4). Let (y, η) be an element of the cosphere bundle $S^*\partial M$ of $T^*\partial M$ and let $\gamma(s) = (y(s), \eta(s))$ be the geodesic with $(y(0), \eta(0)) = (y, \eta)$. Then L^{bf} is given by the union of the leaves $\gamma^2 = \gamma^2(y, \eta)$,

$$\gamma^2 = \mathrm{clos}\Bigg\{\bigg(y, y', \sigma = \frac{x}{x'}, \mu, \mu', \nu, \nu'\bigg) \,\bigg|\, y = y(s),\ y' = y(s'),\ \mu = \eta(s)\sin s,\ \mu' = -\eta(s')\sin s', \\ \nu = -\cos s,\ \nu' = \cos s',\ \sigma = \frac{\sin s}{\sin s'},\ (s, s') \in (0, \pi)^2\Bigg\} \tag{2-6}$$

as (y, η) ranges over $S^*\partial M$. We note that this closure includes the sets

$$T_{\pm} = \{(y, y', \sigma, \mu, \mu', \nu, \nu') \mid y = y',\ \sigma \in \mathbb{R},\ \mu = \mu' = 0,\ \nu = -\nu' = \pm 1\}, \tag{2-7}$$

corresponding to the limits $s, s' \to 0$ and $s, s' \to \pi$.

[3] We denote the new boundary hypersurface of M_{b}^2, created by the blow-up (2-1), by bf. This is slightly at odds with the way bf is used as a boundary hypersurface of $M_{k,b}^2$ — here it really corresponds to taking a section of $M_{k,b}^2$ at fixed $\lambda_* > 0$ — but hopefully no confusion will be caused.

The statement that the spectral measure is a Legendre distribution with respect to the pair of Legendre submanifolds $(L^{\mathrm{bf}}, L^{\sharp})$ means that the Schwartz kernel of the spectral measure can be expressed as an oscillatory function or oscillatory integral, with a phase function that "parametrizes" the Legendre submanifold. We now state what "parametrizes" means, first in the case of a Legendre submanifold L that projects diffeomorphically to the base bf, in the sense that the projection from ${}^{\Phi}T^*_{\mathrm{bf}}M^2_{\mathrm{b}}$ to bf restricts to a (local) diffeomorphism from L to bf. In this case, there exists a function $\Phi : \mathrm{bf} \to \mathbb{R}$ such that (locally) L is the graph of the differential of the function Φ/x or, in coordinates,

$$L=\{\mu=d_y\Phi(y,y',\sigma),\ \mu'=\sigma^{-1}d_{y'}\Phi(y,y',\sigma),\ \nu=\Phi(y,y',\sigma)-\sigma\, d_\sigma\Phi(y,y',\sigma),\ \nu'=d_\sigma\Phi(y,y',\sigma)\}.$$

We say that Φ, or more accurately Φ/x, (locally) *parametrizes* L. In the general case, there always exist (nonunique) functions $\Phi(y, y', \sigma, v)$, depending on extra variables $(v_1, \dots, v_k)$, that locally parametrize L in the sense that

$$L = \bigl\{\mu = d_y\Phi(y,y',\sigma,v),\ \mu'=\sigma^{-1}d_{y'}\Phi(y,y',\sigma,v),\\ \nu=\Phi(y,y',\sigma,v)-\sigma\, d_\sigma\Phi(y,y',\sigma,v),\ \nu'=d_\sigma\Phi(y,y',\sigma,v)\ \big|\ d_v\Phi=0\bigr\}. \tag{2-8}$$

Observe that, if we take the union of the points of (2-6) with $s = s'$, over all $(y, \eta) \in S^*\partial M$, then we get a codimension-1 submanifold of L^{bf}, which is also a codimension-1 submanifold of the conormal bundle of the diagonal $N^*\,\mathrm{diag}_b$, given by

$$N^*\,\mathrm{diag}_b = \{(y, y', \sigma, \mu, \mu', \nu, \nu') \mid y = y',\ \sigma = 1,\ \mu = -\mu',\ \nu = -\nu'\}.$$

Claim. *In a deleted neighbourhood of* $N^*\,\mathrm{diag}_b$, L^{bf} *projects in a* 2:1 *fashion to the base* bf, *i.e.*, $L^{\mathrm{bf}} \setminus N^*\,\mathrm{diag}_b$ *consists of* 2 *sheets*, *each of which projects diffeomorphically to the base* bf, *that are parametrized by the function* $\pm d_{\mathrm{conic}}$, *where* d_{conic} *is the distance function on the cone over* ∂M.

The conic metric d_{conic} has an explicit expression when $d_{\partial M}(y, y') < \pi$. Writing $r = 1/x$ and $r' = 1/x' = \sigma/x$, it takes the form

$$d_{\mathrm{conic}}(y,y',r,r') = \sqrt{r^2+r'^2-2rr'\cos d_{\partial M}(y,y')} = r\sqrt{1+\sigma^2-2\sigma\cos d_{\partial M}(y,y')}. \tag{2-9}$$

Note that $d_{\mathrm{conic}}(y, y', r, r')/r$ indeed has the form $\Phi(y, y', \sigma)/x$ and is smooth provided that $\cos d_{\partial M}(y, y')$ is smooth, i.e., $d_{\partial M}(y, y')$ is less than the injectivity radius on $(\partial M, h)$.

We next explain why we consider the localized (or more precisely microlocalized) spectral measure, by which we mean any of the operators $Q(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q(\lambda)^*$, where $Q(\lambda)$ is a member of our partition of the identity. The reason is, as shown in [Guillarmou et al. 2013b, Section 5], these terms are also Legendre distributions, but associated only to part of the Legendrian, namely to the subset

$$\{(y, y', \sigma, \mu, \mu', \nu, \nu') \in L \mid (y, \mu, \nu), (y', \mu', \nu') \in WF'(Q)\},$$

where $WF'(Q)$ is the support of the symbol[4] of Q. This is localized close to $N^*\,\mathrm{diag}_b \cup T_\pm$ (that is, those points in (2-6) corresponding to $s = s'$) if $WF'(Q)$ is well localized. We can then use the Claim above to

[4]The relevant symbol here is the scattering symbol, or boundary symbol, in the scattering calculus, which is a function on ${}^{\Phi}T^*_{\mathrm{bf}}M^2_{\mathrm{b}}$; see [Melrose 1994].

write this piece of the spectral measure using the conic distance function, except near $N^* \operatorname{diag}_b$ itself, where we can express it as an oscillatory integral using a slightly more complicated form of phase function (as in Proposition 2.6(ii)).

We summarize the information we need from [Guillarmou et al. 2013a; 2013b] concerning the spectral measure:

Proposition 2.6. *Let $Q_j^{\mathrm{low}}(\lambda)$ be a member of the partition of the identity defined above. Let $\eta > 0$ be given. Then, for $j, k = 0, 1$, $Q_j^{\mathrm{low}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_k^{\mathrm{low}}(\lambda)^*$ satisfies the estimates on the right-hand side of* (1-16) *and $Q_j^{\mathrm{low}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j^{\mathrm{low}}(\lambda)^*$, $j \geq 2$, can be written as a finite sum of terms of two types*:

(i) *An oscillatory function of the form*

$$\lambda^{n-1} e^{\pm i\lambda\, d_{\mathrm{conic}}(y, y', 1/x, \sigma/x)} a(y, y', \sigma, x, \lambda), \tag{2-10}$$

where a is supported in $x, x' \leq \eta$ and $d_{\partial M}(y, y') \leq \eta$ and satisfies estimate (1-15).

(ii) *An oscillatory integral of the form*

$$\lambda^{n-1} \int_{\mathbb{R}^{n-1}} e^{i\Phi(y, y', \sigma, v)/\rho} \tilde{a}(y, y', \sigma, v, \rho, \lambda)\, dv, \tag{2-11}$$

where $\tilde{a}$ is smooth in all its arguments and supported in a small neighbourhood of a point $(y_0, y_0, 1, v_0, 0, 0)$ such that $d_v \Phi(y_0, y_0, 1, v_0) = 0$. Moreover, writing $w = (w_1, \dots, w_n)$ for a set of coordinates defining $\operatorname{diag}_b \subset M^2_{k,b}$, *i.e., $w = (y - y', \sigma - 1)$ and $v = (v_2, \dots, v_n)$, one can rotate in the w variables so that the function $\Phi = \Phi(y, w, v)$ has the properties*

$$d_{v_j} \Phi = w_j + O(w_1), \tag{2-12a}$$

$$\Phi = \sum_{j=2}^{n} v_j\, d_{v_j} \Phi + O(w_1), \tag{2-12b}$$

$$d^2_{v_j v_k} \Phi = w_1 A_{jk}, \tag{2-12c}$$

$$\frac{\Phi}{x} = \pm d_{\mathrm{conic}}\Big(y, y', \frac{1}{x}, \frac{\sigma}{x}\Big) \quad \text{if } d_v \Phi = 0, \tag{2-12d}$$

where A_{jk} is nondegenerate for all (y, w, v) in the support of b. Here d_{conic} is as in (2-9).

Proof. The statement about $Q_j^{\mathrm{low}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_k^{\mathrm{low}}(\lambda)^*$ for $j, k = 0, 1$, follows from the microlocal support estimates in [Guillarmou et al. 2013b, Section 5]. In fact, $Q_0^{\mathrm{low}}(\lambda)$ has empty wavefront set, while $Q_1^{\mathrm{low}}(\lambda)$ has wavefront set disjoint from the characteristic variety of $\boldsymbol{H} - \lambda^2$, which contains the microlocal support of $dE_{\sqrt{\boldsymbol{H}}}(\lambda)$. It follows that the operators $Q_j^{\mathrm{low}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_k^{\mathrm{low}}(\lambda)^*$, for $j, k = 0, 1$, vanish rapidly at bf, lb and rb. Also, as shown in [Guillarmou et al. 2013a], $dE_{\sqrt{\boldsymbol{H}}}(\lambda)$ is polyhomogeneous at the other boundary hypersurfaces of $M^2_{k,b}$, namely zf, lb_0, rb_0 and bf_0, vanishing to order $n-1$ at each of these faces. Since the $Q_j^{\mathrm{low}}(\lambda)$ are pseudodifferential operators of order zero, the same is true of the composition $Q_j^{\mathrm{low}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_k^{\mathrm{low}}(\lambda)^*$ for $j, k = 0, 1$ (see [Guillarmou et al. 2013b, Lemma 5.2]). To translate this into an estimate, we observe that λ is a product of boundary defining functions for zf, lb_0, rb_0

and bf_0, while a product of boundary defining functions for bf, lb and rb is $O((1+\lambda d(z,z'))^{-1})$. The estimate (1-16) follows directly.

We next discuss (i) and (ii). Everything in this statement has been proved in [Guillarmou et al. 2013b, Lemma 6.5 and Proposition 6.2] except for the statement that Φ is given by the conic distance function when $d_v\Phi = 0$. To see this, we use the explicit formula (2-9) for the conic distance function, the relation (2-8) and the description of the Legendre submanifold L^{bf} in (2-6). From (2-8), it follows that $\Phi = \nu + \sigma\nu'$. Writing ν and ν' in terms of s and s', using (2-6), we see that

$$d_v\Phi = 0 \implies \Phi = -\cos s + \sigma \cos s'.$$

If we square this then we get

$$d_v\Phi = 0 \implies \Phi^2 = \cos^2 s + \sigma^2 \cos^2 s' - 2\sigma \cos s \cos s'.$$

We can write the right-hand side in the form

$$1 - \sin^2 s + \sigma^2(1 - \sin^2 s') - 2\sigma(\cos(s-s') - \sin s \sin s').$$

Noting that $\sin^2 s + \sigma^2 \sin^2 s' = 2\sigma \sin s \sin s'$, using the expression for σ in (2-6), we see that

$$d_v\Phi = 0 \implies \Phi^2 = 1 + \sigma^2 - 2\sigma \cos d_{\partial M}(y, y'). \qquad \square$$

Remark 2.7. It might help to give an example to show how (2-12) works. In Euclidean space, the Schwartz kernel of the spectral measure $dE_{\sqrt{\Delta}}(\lambda)$ of $\sqrt{\Delta}$ is given by

$$dE_{\sqrt{\Delta}}(\lambda; z, z') = \frac{\lambda^{n-1}}{(2\pi)^n} \int_{\mathbb{S}^{n-1}} e^{i\lambda(z-z')\cdot\zeta}\, d\zeta$$

and one can find the phase function $(z-z')\cdot\zeta$, where $\zeta \in \mathbb{S}^{n-1}$. Locally near $\zeta = (1, 0, \dots, 0)$, we can write $\zeta = (\sqrt{1-|v|^2}, v_2, \dots, v_n)$. Write $x = |z|^{-1}$ and $w = (z-z')/|z|$. Then the phase function becomes

$$\Phi = w_1\sqrt{1 - v_2^2 - \cdots - v_n^2} + \sum_{j=2}^{n} w_j v_j,$$

and we can check that properties (2-12) hold in this case.

3. Spectral measure and partition of the identity at high energies

In the previous section we recalled the partition of the identity operator and the structure of the localized spectral measure for low energy, i.e., $0 < \lambda \le \lambda_0$. We now do the same for high energies, $\lambda \in [\lambda_0, \infty)$. For the sake of convenience, we introduce the semiclassical parameter $h = \lambda^{-1}$ (which should not be confused with h in the metric g), so that we pay our attention to the range $h \in (0, h_0]$, where $h_0 = \lambda_0^{-1}$. The spectral measure of the operator $\boldsymbol{H}$ for high energy was constructed in [Hassell and Wunsch 2008] on the high energy space $\boldsymbol{X}$. Our main task is to adapt each of the main results in the previous section to the high energy setting.

3A. ***High energy space.*** The high energy $\boldsymbol{X}$, introduced in [Hassell and Wunsch 2008], is defined by $\boldsymbol{X} = [0, h_0] \times M_{\mathrm{b}}^2$, where $M_{\mathrm{b}}^2 = [M^2; \partial M \times \partial M]$ is as in (2-1). We label the boundary hypersurfaces in $\boldsymbol{X}$ by rb, lb, bf and mf, according as they are the lifts to $\boldsymbol{X}$ of the faces

$$[0, h_0] \times M \times \partial M, \quad [0, h_0] \times \partial M \times M, \quad [0, h_0] \times \partial M \times \partial M \quad \text{or} \quad \{0\} \times M^2$$

of $[0, h_0] \times M^2$, respectively. The labelling of boundary hypersurfaces is consistent with the notations defined in the low energy space, since when $\lambda \in (C^{-1}, C)$ (where $\lambda = 1/h$) the spaces both have the form $(C^{-1}, C) \times M_{\mathrm{b}}^2$. Recall $\sigma = x/x'$; we can use coordinates (y, y', σ, x', h) near bf and away from rb, and coordinates $(y, y', \sigma^{-1}, x, h)$ near bf and away from lb. We use coordinates (z, z', h) away from bf, rb and lb.

3B. ***Semiclassical scattering pseudodifferential operators.*** We recall the space $\Psi_{\mathrm{sc},h}^{m,l,k}(M; {}^{s\Phi}\Omega^{1/2})$ of semiclassical scattering pseudodifferential operators, introduced by Wunsch and Zworski [2000] based on Melrose's scattering calculus [1994]. Such operators are indexed by the differential order m, the boundary order l and the semiclassical order k. One can express this space in terms of the space with $l = k = 0$ by

$$\Psi_{\mathrm{sc},h}^{m,l,k}(M; {}^{s\Phi}\Omega^{1/2}) = x^l h^{-k} \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2}).$$

The Schwartz kernel of semiclassical pseudodifferential operator $A \in \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2})$ takes the following form on $\boldsymbol{X}$: near the diagonal $\mathrm{diag}_b \subset M_{\mathrm{b}}^2$ and away from bf, it takes the form

$$h^{-n} \int e^{i(z-z')\cdot\zeta/h} a(z, \zeta, h)\, d\zeta\, |dg\, dg'|^{1/2}, \quad n = \dim M, \tag{3-1}$$

while near the boundary of the diagonal, $\mathrm{diag}_b \cap \mathrm{bf}$, it takes the form

$$h^{-n} \int e^{i((y-y')\cdot\mu + (\sigma-1)\nu)/(hx)} a(x, y, \mu, \nu, h)\, d\mu\, d\nu\, |dg\, dg'|^{1/2} \tag{3-2}$$

Here, a is a symbol of order m in the variable ζ or (η, ν) variables and is smooth in the remaining variables. Finally, away from diag_b, the kernel of A is smooth and vanishes to all orders at bf, lb, rb and mf.

Lemma 3.1. *If $A \in \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2})$ then $(h\, \partial_h)^N A$ is also a pseudodifferential operator of order m, i.e., $(h\, \partial_h)^N A \in \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2})$.*

Proof. Away from the diagonal, the result is trivial, as the kernel is smooth and $O(h^\infty)$. So, consider the representations (3-1)–(3-2). The proof is parallel to the argument in Lemma 2.3. By induction, we only need to consider $N = 1$. If $h\partial_h$ hits the function a in (3-2), then a is still a symbol of order m in the (η, ν) variables, smooth in (h, x, y) and supported in $xh \le c$. On the other hand, if $h\, \partial_h$ hits the phase, this is the same as $\nu\partial_\nu + \eta \cdot \partial_\eta$ hitting the phase, as it brings a factor which is homogeneous of degree -1 in h and degree 1 in (ν, η). Integrating by parts, we obtain another symbol $\tilde{a}$ of order m. The argument for (3-1) is analogous. This completes the proof. □

Lemma 3.2. *If $A = A(z, z', h)\, |dg\, dg'|^{1/2} \in \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2})$ and $m < -n$, then A satisfies the kernel bound* (2-5) (*with* $\lambda = h^{-1}$) *for any $N \in \mathbb{N}$.*

Proof. This estimate is straightforward away from the diagonal, as the Schwartz kernel of A vanishes rapidly at all boundaries away from the diagonal. This follows from the nonvanishing of the differential of the phase away from the diagonal. On the other hand, the right-hand side is a positive multiple of $h^{N-n}\rho_{\mathrm{lb}}^N\rho_{\mathrm{bf}}^N\rho_{\mathrm{rb}}^N$ away from the diagonal.

Near the diagonal, we have the representations (3-1)–(3-2). The argument in the case (3-2) is the same as in Lemma 2.4. In the interior case (3-1) we note that the differential operator

$$\frac{1+\Delta_\zeta}{1+h^{-2}|z-z'|^2}$$

leaves the exponential in (3-2) invariant. Applying this differential operator N times and integrating by parts, we see that the integral is bounded by

$$C_N(1+h^{-2}|z-z'|^2)^{-N}$$

for any N. In the interior, the square of the Riemannian distance on M is comparable to $|z-z'|^2$, so the integral is bounded by $C_N(1+h^{-1}d(z,z'))^{-N}$ for any N. □

Corollary 3.3. *If $A \in \Psi_{\mathrm{sc},h}^{m,0,0}(M; {}^{s\Phi}\Omega^{1/2})$ and $m < -n$, then A is bounded $L^2(M^\circ) \to L^2(M^\circ)$ uniformly as $h \to 0$. The same is true for $(h\partial_h)^N A$ for any N.*

Proof. This follows from the kernel bound (2-5) and Schur's test, since there is a uniform volume estimate $cr^n \le V(z,r) \le Cr^n$ for the volume $V(z,r)$ of the ball of radius r centred at $z \in M^\circ$. □

3C. ***High energy partition of the identity.*** We now describe the partition of the identity used in Proposition 1.5 for high energies. Similar to before, these operators are obtained by quantizing symbols which form a partition of unity (independent of h) in the scattering cotangent bundle ${}^{\mathrm{sc}}T^*M$. We modify the open cover $G_0, \dots, G_{N_l}$ from Section 2D by replacing G_0 by a smaller set $\widetilde{G}_0$ given by the points satisfying $x > \epsilon$ and $|\zeta|_g^2 \le \frac{1}{2}$ or $|\zeta|_g^2 \ge \frac{3}{2}$, i.e., the set $\widetilde{G}_0$ is disjoint from the characteristic variety. Then we cover the compact set $\{x \ge \epsilon, |\zeta|_g^2 \in [\frac{1}{4}, 2]\}$, which contains $G_0 \setminus \widetilde{G}_0$, by a finite number $G_{N_l+1}, \dots, G_{N_h}$ of open sets of sufficiently small diameter.

As before, we form a partition of unity subordinate to this refined open cover and take these as the principal symbols of operators Q_j^{high} in the class $\Psi_{\mathrm{sc},h}^{0,0,0}(M; {}^{s\Phi}\Omega^{1/2})$ microsupported in G_j (or $\widetilde{G}_0$ in the case $j = 0$). We will assume that $Q_j^{\mathrm{high}}(\lambda) = Q_j^{\mathrm{low}}(\lambda)$ for intermediate energies $\lambda \sim 1$ and $1 \le j \le N_l$.

3D. ***Localized spectral measure.*** Hassell and Wunsch [2008] showed that the spectral measure for the Laplacian on this setting is, for high energy, a Legendre distribution associated to a pair of Legendre submanifolds L and $L^\sharp$. We briefly explain the meaning of this statement. The Legendre submanifold $L^\sharp$ has already been defined in Section 2E; it lives in the contact manifold ${}^\Phi T^*_{\mathrm{bf}} M_{\mathrm{b}}^2$, living over the boundary hypersurface bf. The new Legendre submanifold L encodes the geodesic flow on T^*M°. It is a submanifold of $\mathbb{R} \times {}^\Phi T^* M_{\mathrm{b}}^2$, which has a natural contact form, described as follows. We write α for the contact form on ${}^{\mathrm{sc}}T^*M$ induced by the inclusion of T^*M° into ${}^{\mathrm{sc}}T^*M$, and α and α' for the lift of this contact form to ${}^\Phi T^* M_{\mathrm{b}}^2$ by the left and right projections, respectively. Writing τ for the coordinate on the $\mathbb{R}$-factor in

$\mathbb{R} \times {}^{\Phi}T^*M_b^2$, the contact form on this space takes the form

$$\alpha + \alpha' - d\tau.$$

Then L is given as follows: Let Σ denote the characteristic variety of $h^2\Delta_g - 1$, given in local coordinates by $\{|\zeta|_{g(z)} = 1\}$ in the interior or $\{|\mu|^2_{h(x,y)} + \nu^2 = 1\}$ near the boundary. Then L is given in terms of the geodesic flow G_t by

$$L = \{(q, q', \tau) \mid q, q' \in \Sigma,\ q = G_\tau(q')\} \tag{3-3}$$

(this follows from [Guillarmou et al. 2013b, Equation 7.9] and the discussion following). In $\mathbb{R} \times {}^{\Phi}T^*M_b^2$, L can be restricted to $\mathbb{R} \times {}^{\Phi}T^*_{\mathrm{bf}}M_b^2$, i.e., restricted to lie over bf, then, forgetting the τ component, we obtain the Legendre submanifold L^{bf} from Section 2E.[5]

As in Section 2E, the statement that an operator is Legendrian with respect to L means that its Schwartz kernel can be expressed as an oscillatory function or oscillatory integral using a phase function that locally parametrizes L. In the interior of $\boldsymbol{X}$, this means a function $\Psi(z, z', v)$ such that, locally, using coordinates $(z, \zeta, z', \zeta', \tau)$ on $\mathbb{R} \times {}^{\Phi}T^*M_b^2$, we have

$$L = \{(z, d_z\Psi, z', d_{z'}\Psi, \Psi) \mid d_v\Psi = 0\}.$$

In particular, τ is equal to the value of the phase function when $d_v\Psi = 0$. If there are no v variables, the condition $d_v\Psi = 0$ is omitted and then L is (essentially) the graph of the differential of Ψ. Near the boundary bf, we use local coordinates $(x, y, y', \sigma, \mu, \nu, \mu', \nu', \tau)$ and then a local parametrization of L is given by a function $\Psi(x, y, y', \sigma, v)/x$ such that

$$L = \{(x, y, y', \sigma, d_y\Psi, \Psi - x\, d_x\Psi, -\sigma\, d_\sigma\Psi, \sigma^{-1}\, d_{y'}\Psi, d_\sigma\Psi, \Psi) \mid d_v\Psi = 0\}.$$

We give some consequences of this result for the localized spectral measure needed in this paper. As in the low energy case, the localized spectral measure refers to any operator of the form $Q^{\mathrm{high}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q^{\mathrm{high}}(\lambda)^*$ where $Q^{\mathrm{high}}(\lambda)$ is a member of the partition of the identity operator from Section 3C. As above, we write $h = 1/\lambda$.

Proposition 3.4. *Let $Q_j^{\mathrm{high}}(\lambda)$ be a member of the partition of the identity defined above. Then, for $j, k = 0, 1$, the operator $Q_j^{\mathrm{high}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_k^{\mathrm{high}}(\lambda)^*$ satisfies the estimates on the right-hand side of* (1-16) *and $Q_j^{\mathrm{high}}(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j^{\mathrm{high}}(\lambda)^*$, $j \geq 2$, can be written as a finite sum of terms of the following three types*:

(i) *An oscillatory function of the form*

$$h^{-(n-1)} e^{\pm i d(z,z')/h} \tilde{a}(z, z', h), \tag{3-4}$$

where $\tilde{a}$ satisfies estimate (1-15).

(ii) *An oscillatory integral supported in $x, x' \geq \epsilon$ of the form*

$$h^{-(n-1)} \int_{\mathbb{R}^{n-1}} e^{i\Psi(z,z',v)/h} b(z, z', v, h)\, dv, \tag{3-5}$$

[5] The relation between the various Legendre submanifolds is explained in detail in [Hassell and Wunsch 2008, Part 1].

where b is smooth in all its arguments and supported in a small neighbourhood of a point $(z_0, z_0, v_0, 0)$ *such that* $d_v\Psi(z_0, z_0, v_0) = 0$. *Moreover, writing* $w = z - z'$ *and* $v = (v_2, \dots, v_n)$, *one can rotate in the* w *variables so that the function* $\Psi = \Psi(z, w, v)$ *has the properties*

$$d_{v_j}\Psi = w_j + O(w_1), \tag{3-6a}$$

$$\Psi = \sum_{j=2}^{n} v_j\, d_{v_j}\Psi + O(w_1), \tag{3-6b}$$

$$d^2_{v_j v_k}\Psi = w_1 A_{jk}, \tag{3-6c}$$

$$\Psi(z, z', v) = \pm d(z, z') \quad \text{if } d_v\Psi = 0, \tag{3-6d}$$

where A_{jk} *is nondegenerate at* (z_0, z_0, v_0) *and* $d(z, z')$ *is the Riemannian distance function on* $M^\circ \times M^\circ$.

(iii) *An oscillatory integral supported near* $x = x' = 0$ *of the form*

$$h^{-(n-1)} \int_{\mathbb{R}^{n-1}} e^{i\Psi(y,y',\sigma,x,v)/(hx)} b(y, y', \sigma, x, v, h)\, dv, \tag{3-7}$$

where b is smooth in all its arguments and supported in a small neighbourhood of a point $(y_0, y_0, 1, 0, v_0, 0)$ *such that* $d_v\Psi(y_0, y_0, 1, v_0) = 0$. *Moreover, writing* $w = (w_1, \dots, w_n)$ *for a set of coordinates defining* $\operatorname{diag}_b \subset M_b^2$, *i.e.,* $w = (y - y', \sigma - 1)$ *and* $v = (v_2, \dots, v_n)$, *one can rotate in the* w *variables so that the function* $\Psi = \Psi(y, w, x, v)$ *has the properties*

$$d_{v_j}\Psi = w_j + O(w_1), \tag{3-8a}$$

$$\Psi = \sum_{j=2}^{n} v_j\, d_{v_j}\Psi + O(w_1), \tag{3-8b}$$

$$d^2_{v_j v_k}\Psi = w_1 A_{jk}, \tag{3-8c}$$

$$\Psi/x = \pm d(z, z') \quad \text{if } d_v\Psi = 0, \tag{3-8d}$$

where A_{jk} *is nondegenerate at* $(y_0, y_0, 1, 0, v_0, 0)$.

Remark 3.5. Since $\lambda = 1/h$, this is an analogue of Proposition 2.6 for the case $X = [0, h_0] \times M_b^2$.

Proof. The proof is analogous to the proof of Proposition 2.6, with the main difference being that the computation takes place over the whole of M_b^2 (including the interior), not just at the boundary as in the low energy case. We prove (ii), i.e., we work in the interior of M_b^2, using coordinates (z, z'), with z a coordinate on the left copy of M° and z' on the right copy. The proof for (iii) is only notationally different.

As in the low energy case, the Legendre submanifold L has the property that it intersects $N^* \operatorname{diag}_b$ in a codimension-1 submanifold and, in a deleted neighbourhood of $N^* \operatorname{diag}_b$, it projects in a 2:1 fashion down to the base, $\mathrm{mf} = M_b^2$, so that the two sheets are parametrized by the phase functions $\pm d(z, z')$.

We now apply [Guillarmou et al. 2013b, Lemma 7.6 and (ii) of Lemma 7.7]. This tells us that, for any point in the microlocal support of $Q_j^{\mathrm{high}}(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_j^{\mathrm{high}}(\lambda)^*$, either there is a neighbourhood in which L projects diffeomorphically to the base M_b^2 or the point lies at the conormal bundle to the diagonal, i.e.,

$z = z'$ and $\zeta = -\zeta'$. In the former case, the function $\pm d(z, z')$ can be used directly as the phase function and we obtain the statement (i) in the proposition. In the latter case, a phase function Ψ depending on $n-1$ variables $v_2, \dots, v_n$ can be constructed following the general approach of [Guillarmou et al. 2013b, Proposition 7.5]. Since this was not written down explicitly in the coordinates (z, z') valid in the interior of M_b^2, we sketch briefly how this is done. It follows from the proof of Lemma 7.6 of [Guillarmou et al. 2013b] that we can rotate coordinates so that $w_1, \zeta_2, \dots, \zeta_n, z'$ give coordinates on L locally. (The proof of Lemma 7.6 shows that one can take $(\tau, \zeta_2, \dots, \zeta_n, z')$ but, since it is also shown that $\partial z_1/\partial\tau \neq 0$, one can substitute z_1 for τ and then substitute $w_1 = z_1 - z_1'$ for z_1.) One can therefore express the functions $w_2, \dots, w_n$ and τ on L as smooth functions $W_j(w_1, \zeta_2, \dots, \zeta_n, z')$ and $T(w_1, \zeta_2, \dots, \zeta_n, z')$ of these coordinates. Then the function

$$\Psi(w, z', v) = \sum_{j=2}^{n} (w_j - W_j(w_1, \zeta_2, \dots, \zeta_n, z'))v_j + T(w_1, \zeta_2, \dots, \zeta_n, z')$$

satisfies the requirements of (3-6) and parametrizes L locally. This is shown by adapting the argument of [Guillarmou et al. 2013b, Proof of Proposition 6.2] in a straightforward way (which itself is a minor variation on [Hörmander 1985, Theorem 21.2.18]), so we omit the details. This establishes part (iii) of the proposition. When working close to $x = x' = 0$, we need to use coordinates as in [Guillarmou et al. 2013b, Proposition 7.5] and apply [Guillarmou et al. 2013b, Lemma 7.6 and (i) of Lemma 7.7], and we end up with the statement in part (ii). □

Remark 3.6. The Lagrangian L is smooth up to the boundary when viewed as a submanifold in the "scattering-fibred cotangent bundle" described in [Guillarmou et al. 2013a]. The boundary at bf is naturally isomorphic to L^{bf} in Proposition 2.6. Correspondingly, we find that the distance function $d(z, z')$ on M_b^2 satisfies

$$d(z, z') - d_{\mathrm{conic}}\Big(y, y', \frac{1}{x}, \frac{\sigma}{x}\Big) = e(z, z')$$

is a bounded function on M_b^2 or, more precisely, on that part of M_b^2 where $x, x' \le \eta$ and $d_{\partial M}(y, y') \le \eta$ for sufficiently small η (see [Hassell et al. 2005, Lemma 9.4]). From this we see that the results of Propositions 2.6 and 3.4 are compatible, as the factor $\exp(i\lambda e(z, z'))$ — which is the discrepancy between (2-10) and (3-4) and between (2-12d) and (3-6d) — can be absorbed in the symbols $\tilde{a}$ and b, respectively.

Remark 3.7. The results of this paper could be extended to long-range scattering metrics, as treated in [Hassell et al. 2006]. However, this would require an extension of the results of [Hassell and Vasy 2001; Hassell and Wunsch 2008; Guillarmou et al. 2013a] to Lagrangian submanifolds which are only conormal, rather than smooth, at the boundary. If this were done, then the discrepancy $e(z, z')$ between the distance function and the conic distance function would no longer be smooth or even bounded, but rather conormal at the boundary with a bound of the form $(x + x')^{-1+\epsilon}$ at the boundary of M_b^2, i.e., a bit smaller than the distance functions themselves. In this case, the correct description of the localized spectral measure would be with the true distance function $d(z, z')$ as phase function, rather than (2-10), which is only true in the short-range case.

The assumption on the potential could also be weakened; for example, one could assume that V only decays as $x^{2+\epsilon}$ as $x \to 0$, and is only conormal, rather than smooth, as $x \to 0$, instead of (1-5). However, if one assumes only $O(x^2)$ decay then it is not clear whether Theorem 1.1 will hold. For example, if $V \in x^2 C^\infty(M)$ and $V_0 := x^{-2} V|_{\partial M}$ takes values in the range $\left(-\frac{1}{4}(n-2)^2, 0\right)$, then it follows from [Guillarmou et al. 2013a, Corollary 1.5] that the $L^1 \to L^\infty$ norm of the propagator is at least a constant times $t^{-(\nu_0+1)}$ as $t \to \infty$, where ν_0^2 is the smallest eigenvalue of $\Delta_{\partial M} + V_0 + \frac{1}{4}(n-2)^2$. Under the above assumption on the range of V_0, we see that $\nu_0 < \frac{1}{2}n - 1$. This implies that the *dispersive* estimate (1-12) will no longer be valid as $|t-s| \to \infty$. However, the implications of that for the global-in-time Strichartz estimates are not clear; in the case of inverse-square potentials on $\mathbb{R}^n$, global-in-time Strichartz estimates hold despite the fact that the dispersive estimate is not known to hold for negative inverse-square potentials [Burq et al. 2004b] (for positive inverse-square potentials, the dispersive estimate is proved in [Fanelli et al. 2013]).

The problem, however, is only with the *long*-time Strichartz estimates; for estimates on a finite time interval, the decay condition on V as $x \to 0$ could be weakened considerably.

4. Proof of Proposition 1.5

We now prove Proposition 1.5. We define our partition of unity Q_j by combining the low-energy and high-energy partitions. We choose a cutoff function $\chi(\lambda)$ supported in $[0, 2]$ such that $1-\chi$ is supported in $[1, \infty)$ and define

$$\begin{aligned} Q_1(\lambda) &= \chi(\lambda)(Q_0^{\mathrm{low}} + Q_1^{\mathrm{low}}) + (1-\chi(\lambda))(Q_0^{\mathrm{high}} + Q_1^{\mathrm{high}}), \\ Q_j(\lambda) &= \chi(\lambda) Q_j^{\mathrm{low}} + (1-\chi(\lambda)) Q_j^{\mathrm{high}} \qquad \text{for } 2 \le j \le N_l, \\ Q_j(\lambda) &= (1-\chi(\lambda)) Q_j^{\mathrm{high}} \qquad \text{for } N_l + 1 \le j \le N. \end{aligned} \tag{4-1}$$

We first note that the term with $Q_1(\lambda)$ satisfies (1-14) (with only the "b" term present) and (1-16), according to Propositions 2.6 and 3.4. (In the case of low energies we also need to use Remark 3.6, which tells us that we can replace the distance function by the conic distance function d_{conic} in (1-14) without affecting the estimates on the amplitudes $a_\pm$.)

Next we prove the proposition for low energies, i.e., for $\lambda \le 2$, and $j \ge 2$. Consider the second type of representation, (2-11), in Proposition 2.6. We break the estimate into various cases. We first observe that estimates of the form (1-15) and (1-16) are unaffected by multiplication by a cutoff function of the form $\chi(\lambda d(z, z'))$, where $\chi \in C_c^\infty(\mathbb{R})$. Therefore, we may treat the cases $\lambda d(z, z') \lesssim 1$ and $\lambda d(z, z') \gtrsim 1$ separately. Consider first the case $\lambda d(z, z') \lesssim 1$ or, equivalently, $|w| \lesssim \rho$. In this case, we show that (2-11) has the form (1-14), where only the "b" term is present, satisfying (1-16). Thus, we need to show that

$$(\lambda\, \partial_\lambda)^\alpha \int_{\mathbb{R}^{n-1}} e^{i\lambda\Phi(y,w,v)/x} \tilde{a}\Big(\lambda, \frac{x}{\lambda}, y, w_1, v\Big)\, dv$$

is uniformly bounded. For $\alpha = 0$ this is obvious. So consider the effect of applying $\lambda\, \partial_\lambda$. This is harmless when it hits $\tilde{a}$. When it hits the phase, it brings down a factor $i\lambda\Phi/x$. We have $\lambda\Phi/x = \Phi/\rho = v \cdot d_v\Phi/\rho + O(w_1/\rho)$ and, since $|w| \lesssim \rho$, the $O(w_1/\rho)$ is harmless. To treat the $v \cdot d_v\Phi/\rho$ term, we can write, using (2-12b),

$$\frac{v \cdot d_v \Phi}{\rho} e^{i\Phi/\rho} = -iv \cdot d_v e^{i\Phi/\rho},$$

and integrating by parts we see that this term is $O(1)$ after integration. Repeated applications of $\lambda\partial_\lambda$ are treated similarly.

Second, suppose that $|w| \geq C\rho$ for some large C but that $|w_1| \leq \rho$. For large enough C, this means that $d_{v_j}\Phi \neq 0$ for some $j \geq 2$ since, by (2-12a), we have $d_{v_j}\Phi = w_j - O(w_1)$. So, by choosing j so that $|w_j|$ is maximal and then C large enough, we have $|d_{v_j}\Phi| \geq c|w|$. Then we can write

$$e^{i\Phi/\rho} = \left(\frac{\rho\, d_{v_j}}{i\, d_{v_j}\Phi}\right)^N e^{i\Phi/\rho}$$

and integrate by parts. Each integration by parts gains us a factor of $\rho/|w|$. Thus we can estimate (2-11) by $(1+|w|/\rho)^{-K} = (1+\lambda d(z,z'))^{-K}$ for any K. Estimating the terms for $\alpha > 0$ is done just as in the first case above.

Third, suppose that $|w| \geq C|w_1|$ for some large C and that $|w_1| \geq \rho$. Then we can integrate by parts and gain any number of factors of $(1+\lambda d(z,z'))^{-1}$ as in the second case above.

Finally we come to the case where $|w_1| \geq \rho$ and $|w_1|$ is comparable to $|w|$. In this case, we have removed a neighbourhood of $N^*\operatorname{diag}_b$ from the microlocal support of the localized spectral measure. As discussed in Section 2, in this region the Lagrangian L^{bf} is a union of two sheets, each of which projects diffeomorphically to the base bf and is parametrized by the phase function $\pm d_{\mathrm{conic}}$ (in terms of the phase function Φ as in (2-11)–(2-12), this simply corresponds to the sign of w_1). We can thus split this case into two parts, according to the sign of w_1, which give rise to the "$\pm$" terms in (1-14).

In this case, the key is to exploit property (2-12c). Define

$$\widetilde{\Phi}(x,y,w,v) = |w_1|^{-1}(\Phi(y,w,v) \mp x d(z,z')) \tag{4-2}$$

and let $\omega = |w_1|/\rho$; then we need to estimate

$$\lambda^\alpha\, \partial_\lambda^\alpha a(\lambda,z,z') = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!}\omega^\beta \int_{\mathbb{R}^{n-1}} e^{i\omega\widetilde{\Phi}(x,y,w,v)}\widetilde{\Phi}^\beta(\lambda^\gamma\, \partial_\lambda^\gamma \tilde{a})(\lambda,\rho,y,w_1,v)\,dv.$$

Let $\tilde{b} = \lambda^\gamma\, \partial_\lambda^\gamma \tilde{a}$; then $|\partial_\lambda^\gamma \tilde{b}| \leq C_\gamma \lambda^{-\gamma}$. Thus, noting $\omega \geq 1$, it suffices to show that, for any $0 \leq \beta \leq \alpha$,

$$\left|\int_{\mathbb{R}^{n-1}} e^{i\omega\widetilde{\Phi}(x,y,w,v)}(\omega\widetilde{\Phi})^\beta \tilde{b}(\lambda,\rho,y,w_1,v)\,dv\right| \leq C\omega^{-(n-1)/2}. \tag{4-3}$$

To proceed, we fix (x,y,w) with $w \neq 0$ (and hence $w_1 \neq 0$ due to our assumption that $|w_1|$ is comparable to $|w|$). We use a cutoff function Υ to divide the v integral into two parts: one on the support of Υ, in which $|d_v\widetilde{\Phi}| \geq \frac{1}{2}\tilde{\epsilon}$, and the other on the support of $1-\Upsilon$, in which $|d_v\widetilde{\Phi}| \leq \tilde{\epsilon}$. On the support of Υ, we integrate by parts in v and gain any power of ω^{-1}, proving (4-3). On the support of $1-\Upsilon$, we make the variable change

$$(v_2,\ldots,v_n) \to (\theta_2,\ldots,\theta_n), \quad \theta_i = d_{v_i}\widetilde{\Phi}, \quad i = 2,\ldots,n.$$

Note that, by property (2-12c),

$$\frac{\partial\theta_j}{\partial v_k} = d^2_{v_j v_k}\widetilde{\Phi} = \pm A_{jk}.$$

The nondegeneracy of A_{jk} shows that this change of variables is locally nonsingular provided $\tilde{\epsilon}$ is sufficiently small. Thus, for each point v in the support of $1-\Upsilon$, there is a neighbourhood in which we can change variables to θ as above. Using the compactness of the support of b in (2-11), we see that there are a finite number of neighbourhoods covering the intersection of the support of Υ and the v-support of b. For simplicity of exposition, we assume that there is only one such neighbourhood U below.

Let $\mathcal{B}_\delta := \{\theta : |\theta| \le \delta\}$ and choose a C^∞ function $\chi_{\mathcal{B}_\delta}(\theta)$ which equals 1 on the set $\mathcal{B}_\delta$ but equals 0 outside $\mathcal{B}_{2\delta}$, with bounds on the derivatives given by

$$|\nabla^{(j)}_\theta \chi_{\mathcal{B}_\delta}(\theta)| \le C\delta^{-j}.$$

Here δ is a parameter to be chosen later (depending on ω). Consider the integral (4-3) after changing variables and with the cutoff function $\chi_{\mathcal{B}_\delta}(\theta)$ inserted (note that $1-\Upsilon = 1$ on the support of $\chi_{\mathcal{B}_\delta}(\theta)$, provided $\delta \le \frac{1}{2}\tilde{\epsilon}$):

$$\left|\int e^{i\omega\widetilde{\Phi}(x,y,w,\theta)}(\omega\widetilde{\Phi})^\beta \tilde{b}(\lambda,\rho,y,w_1,\theta)\chi_{\mathcal{B}_\delta}(\theta)\frac{d\theta}{|A^{-1}(y,w,\theta)|}\right|.$$

Using property (2-12d), we see that $\widetilde{\Phi} = 0$ when $\theta = 0$. Also, due to our choice of θ, we have $d_\theta\widetilde{\Phi} = 0$ when $\theta = 0$, so $\widetilde{\Phi} = O(|\theta|^2)$. Hence,

$$\left|\omega^\beta\int e^{i\omega\widetilde{\Phi}(x,y,w,\theta)}\widetilde{\Phi}^\beta \tilde{b}(\lambda,\rho,y,w_1,\theta)\chi_{\mathcal{B}_\delta}(\theta)\frac{d\theta}{|A^{-1}(y,w,\theta)|}\right| \le C(\omega\delta^2)^\beta\delta^{n-1}.$$

It remains to treat the integral with cutoff $1-\chi_{\mathcal{B}_\delta}(\theta)$ inserted. Notice that $|d_\theta\widetilde{\Phi}|$ is comparable to $|\theta|$ since $d_\theta\widetilde{\Phi} = 0$ when $\theta = 0$, and

$$d^2_{\theta_i\theta_j}\widetilde{\Phi} = \sum_{k,l}(A^{-1})_{il}(A^{-1})_{jk}\, d^2_{v_k v_l}\widetilde{\Phi}$$

is nondegenerate when $\theta = 0$. We define the differential operator L by

$$L = \frac{-i\, d_\theta\widetilde{\Phi}\cdot\partial_\theta}{\omega|d_\theta\widetilde{\Phi}|^2}.$$

Then the adjoint operator is given by

$${}^tL = -L + \frac{i}{\omega}\left(\frac{\Delta_\theta\widetilde{\Phi}}{|d_\theta\widetilde{\Phi}|^2} - 2\frac{d^2_{\theta_j\theta_k}\widetilde{\Phi}\, d_{\theta_j}\widetilde{\Phi}\, d_{\theta_k}\widetilde{\Phi}}{|d_\theta\widetilde{\Phi}|^4}\right).$$

Since $Le^{i\omega\widetilde{\Phi}} = e^{i\omega\widetilde{\Phi}}$, we integrate by parts N times to obtain

$$\left|\int e^{i\omega\widetilde{\Phi}(x,y,w,\theta)}(\omega\widetilde{\Phi})^\beta \tilde{b}(\lambda,\rho,y,w_1,\theta)(1-\chi_{\mathcal{B}_\delta}(\theta))(1-\Upsilon)\, d\theta\right|$$
$$\le C\int\left|({}^tL)^N\big((\omega\widetilde{\Phi})^\beta \tilde{b}(\lambda,\rho,y,w_1,\theta)(1-\chi_{\mathcal{B}_\delta}(\theta))(1-\Upsilon)\big)\right| d\theta.$$

Inductively, we find that

$$\left|({}^tL)^N\big((\omega\widetilde{\Phi})^\beta\tilde{b}(1-\chi_{\mathscr{B}_\delta})(1-\Upsilon)\big)\right| \le C\omega^{-N+\beta}\max\{|\theta|^{2\beta-2N}, |\theta|^{2\beta-N}\delta^{-N}\}.$$

Choosing N large enough, we get

$$\begin{aligned}\left|\int e^{i\omega\widetilde{\Phi}(x,y,w,\theta)}(\omega\widetilde{\Phi})^\beta\tilde{b}(\lambda,\rho,y,w_1,\theta)(1-\chi_{\mathscr{B}_\delta})(1-\Upsilon)\,d\theta\right| &\le \omega^{-N+\beta}\int_{|\theta|\ge\delta}(|\theta|^{2\beta-2N}+|\theta|^{2\beta-N}\delta^{-N})\,d\theta\\ &\le C\omega^{-N+\beta}\delta^{2\beta-2N}\delta^{n-1}.\end{aligned}$$

Choose $\delta=\omega^{-1/2}$ to balance the two parts of the integral (with $\chi_{\mathscr{B}_\delta}$ and with $1-\chi_{\mathscr{B}_\delta}$). We finally obtain

$$\left|\int e^{i\omega\widetilde{\Phi}(x,y,w,\theta)}(\omega\widetilde{\Phi})^\beta\tilde{b}(\lambda,\rho,y,w_1,\theta)(1-\Upsilon)\,d\theta\right| \le C\omega^{-(n-1)/2},$$

which proves (4-3), as desired.

We next sketch how to prove (1-16) in the high-energy case $i > N_l$. In terms of Proposition 3.4, consider a term of type (iii); it suffices to show

$$a(h,z,z') = e^{\mp id(z,z')/h}\int_{\mathbb{R}^{n-1}} e^{i\Psi(y,w,x,v)/(xh)}b(h,x,y,w_1,v)\,dv,$$

satisfies

$$|(h\,\partial_h)^\alpha a(h,z,z')| \le C_\alpha\left(1+\frac{|w|}{xh}\right)^{-\frac{n-1}{2}}.$$

Notice that $\lambda=1/h$ and Ψ has the same properties (2-12a)–(2-12d) as Φ. Therefore the low energy proof works verbatim, with the argument x of Ψ acting as a smooth parameter, and leads to the desired conclusion. The proof in case (ii) works in exactly the same way, with w given by $z-z'$.

Remark 4.1. To illustrate this theorem, consider the case of the spectral measure on flat $\mathbb{R}^3$, which is

$$dE_{\sqrt{\Delta}}(\lambda)(z,z') = \frac{1}{2\pi^2}\frac{\lambda^2\sin\lambda|z-z'|}{\lambda|z-z'|}\,d\lambda.$$

We decompose this, using the cutoff function χ as in (4-1), according to the size of $\lambda|z-z'|$. Where $\lambda|z-z'|\ge 1$, that is, more than one wavelength from the diagonal, we split the sine factor into exponential terms. Within $O(1)$ wavelengths of the diagonal, however, we keep the sine factor as is, to exploit the cancellation in the difference $e^{+i\lambda|z-z'|}-e^{-i\lambda|z-z'|}$ when $\lambda|z-z'|$ is small. This gives us an expression

$$\frac{\lambda^2}{2\pi^2}\left((1-\chi)(\lambda|z-z'|)\frac{e^{i\lambda|z-z'|}}{2i\lambda|z-z'|}-(1-\chi)(\lambda|z-z'|)\frac{e^{-i\lambda|z-z'|}}{2i\lambda|z-z'|}+\chi(\lambda|z-z'|)\frac{\sin\lambda|z-z'|}{\lambda|z-z'|}\right).$$

This is a decomposition into "$\pm$" and "b" terms as in (1-14), where the amplitudes satisfy (1-15) and (1-16). So, we can think of the b term as the near-diagonal term and the other terms as related to the two sheets of the Lagrangian L or L^{bf}, which are separated away from the diagonal. The function of the microlocalizing operators $Q_j(\lambda)$ (which are not required in the case of flat Euclidean space) is to remove parts of the Lagrangian that do not project diffeomorphically to the base.

5. L^2 estimates

In this section, we prove $L^2 \to L^2$ estimates on microlocalized versions of the Schrödinger propagator, using the operator partition of unity Q_j described at the beginning of the previous section, based on [Guillarmou et al. 2013b].

We begin by defining microlocalized propagators. First we give a formal definition. It is not immediately clear that the formal definition is well defined, so our first task is to show this. We do so by showing that each microlocalized propagator is a bounded operator on L^2. This serves both to show the well-definedness of each microlocalized propagator and to establish the $L^2 \to L^2$ estimate needed for the abstract Keel–Tao argument.

We define, as in the introduction,

$$U_j(t) = \int_0^\infty e^{it\lambda^2} Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda), \tag{5-1}$$

where Q_j is the decomposition defined in (4-1).

Our first task is to make sense of this expression. We do this by showing that each $U_j(t)$ is a bounded operator on $L^2(M^\circ)$. We have:

Proposition 5.1. *For each j, the integral* (5-1) *defining $U_j(t)$ is well defined on each finite interval and converges on $\mathbb{R}_+$ in the strong operator topology to define a bounded operator on $L^2(M^\circ)$. Moreover, the operator norm of $U_j(t)$ on $L^2(M^\circ)$ is bounded uniformly for $t \in \mathbb{R}$. Finally, we have*

$$\sum_j U_j(t) = e^{it\boldsymbol{H}}. \tag{5-2}$$

Proof. Suppose that $A(\lambda)$ is a family of bounded operators on $L^2(M^\circ)$, compactly supported and C^1 in $\lambda \in (0, \infty)$. Integrating by parts,

$$\int_0^\infty A(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)$$

is given by

$$-\int_0^\infty \Big(\frac{d}{d\lambda} A(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda)\, d\lambda.$$

In view of Corollaries 2.5 and 3.3, we can take $A(\lambda)$ to be a smooth function of λ with compact support in $(0, \infty)$ multiplied by $e^{it\lambda^2} Q_j(\lambda)$. This means that the integral (5-1) is well defined over any compact interval in $(0, \infty)$. We need to show that the integral over the whole of $\mathbb{R}_+$ converges in the strong operator topology. To do so, we introduce a dyadic partition of unity on the positive λ axis by choosing $\phi \in C_c^\infty\big(\big[\tfrac{1}{2}, 2\big]\big)$, taking values in [0, 1], such that

$$\sum_{m \in \mathbb{Z}} \phi\Big(\frac{\lambda}{2^m}\Big) = 1.$$

We now define

$$U_{j,m}(t) = -\int_0^\infty \frac{d}{d\lambda}\Big(e^{it\lambda^2} \phi\Big(\frac{\lambda}{2^m}\Big) Q_j(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda). \tag{5-3}$$

We next show that the sum over m of the operators $U_{j,m}(t)$ in (5-3) is well defined. For this we use the Cotlar–Stein lemma, which we recall here (we use the version in [Grafakos 2009, Chapter 8]):

Lemma 5.2 (Cotlar–Stein lemma). *Suppose that $\{A_j\}$ is a sequence of bounded linear operators on a Hilbert space H such that*

$$\|A_m^* A_n\|_{H\to H} \le (\gamma(m-n))^2, \quad \|A_m A_n^*\|_{H\to H} \le (\gamma(m-n))^2, \tag{5-4}$$

where $\{\gamma(m)\}_{m\in\mathbb{Z}}$ is a sequence of positive constants such that $C = \sum_{m\in\mathbb{Z}} \gamma(m) < \infty$. Then, for all $f \in H$, the sequence $\sum_{|m|\le N} A_m f$ converges as $N \to \infty$ to an element $Af \in H$. The operators $A = \sum_m A_m$ and $A^* = \sum_m A_m^*$ so defined (in the strong operator topology) satisfy

$$\|A\|_{H\to H} \le C, \quad \|A^*\|_{H\to H} \le C. \tag{5-5}$$

Moreover, the operator norms of $\sum_{m\in J} A_m$ and $\sum_{m\in J} A_m^*$ are bounded by C for any finite subset J of the integers.

We also use the following lemma:

Lemma 5.3. *Suppose that $A_l(\lambda)$ for $l = 1, 2$ is a family of operators compactly supported in λ in the open interval $(0,\infty)$ with $A_l(\lambda)$ and $\partial_\lambda A_l(\lambda)$ uniformly bounded on $L^2(M^\circ)$. Define*

$$B_l = \int A_l(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda).$$

Then

$$B_1 B_2^* = \int A_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, A_2(\lambda)^*,$$

where by definition the last expression is equal to

$$\int \Big(-\frac{d}{d\lambda} A_1(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda) A_2(\lambda) - A_1(\lambda) E_{\sqrt{\boldsymbol{H}}}(\lambda) \Big(\frac{d}{d\lambda} A_2(\lambda)\Big). \tag{5-6}$$

Proof. We compute

$$\begin{aligned}
B_1 B_2^* &= \iint \Big(\frac{d}{d\lambda} A_1(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda) E_{\sqrt{\boldsymbol{H}}}(\mu) \Big(\frac{d}{d\mu} A_2(\mu)^*\Big)\, d\lambda\, d\mu \\
&= \iint_{\lambda\le\mu} \Big(\frac{d}{d\lambda} A_1(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda) \Big(\frac{d}{d\mu} A_2(\mu)^*\Big)\, d\lambda\, d\mu \\
&\quad + \iint_{\mu\le\lambda} \Big(\frac{d}{d\lambda} A_1(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\mu) \Big(\frac{d}{d\mu} A_2(\mu)^*\Big)\, d\lambda\, d\mu \\
&= \int \Big(\frac{d}{d\lambda} A_1(\lambda)\Big) E_{\sqrt{\boldsymbol{H}}}(\lambda)(-A_2(\lambda)^*)\, d\lambda + \int (-A_1(\mu)) E_{\sqrt{\boldsymbol{H}}}(\mu) \Big(\frac{d}{d\mu} A_2(\mu)^*\Big)\, d\mu \\
&= (5\text{-}6).
\end{aligned} \tag{5-7}$$

This concludes the proof. □

Now we show that the sum in (5-3) is well defined. We first note a simplification: since the $Q_j(\lambda)$ are a partition of the identity, we have

$$V_m(t) := \sum_{j=1}^{N} U_{j,m}(t) = \int e^{it\lambda^2}\chi(\lambda)\phi\left(\frac{\lambda}{2^m}\right) dE_{\sqrt{H}}(\lambda),$$

which is clearly bounded on $L^2(M^\circ)$ with operator norm bounded by 1 using spectral theory. Moreover, the sum of any subset of the V_m converges strongly to an operator with norm bounded by 1. Due to this, we may ignore the case $j=1$ and prove the L^2 boundedness only for $j \geq 2$.

We have, by Lemma 5.3,

$$\begin{aligned} U_{j,m}(t)U_{j,n}(t)^* &= \int \chi(\lambda)^2\phi\left(\frac{\lambda}{2^m}\right)\phi\left(\frac{\lambda}{2^n}\right) Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_j(\lambda)^* \\ &= -\int \frac{d}{d\lambda}\left(\chi(\lambda)^2\phi\left(\frac{\lambda}{2^m}\right)\phi\left(\frac{\lambda}{2^n}\right) Q_j(\lambda)\right) E_{\sqrt{H}}(\lambda) Q_j(\lambda)^* \\ &\qquad - \int \chi(\lambda)^2\phi\left(\frac{\lambda}{2^m}\right)\phi\left(\frac{\lambda}{2^n}\right) Q_j(\lambda) E_{\sqrt{H}}(\lambda) \frac{d}{d\lambda} Q_j(\lambda)^*. \end{aligned} \tag{5-8}$$

We observe that this is independent of t and is identically zero unless $|m-n| \leq 2$. When $|m-n| \leq 2$, we note that the integrand is a bounded operator on L^2, with an operator bound of the form C/λ, where C is uniform, as we see from Corollary 2.5 and the support property of ϕ. The integral is therefore uniformly bounded, as we are integrating over a dyadic interval in λ.

We next consider the operators $U^*_{j,m}(0)U_{j,n}(0)$, just in the case $t=0$. This has an expression

$$\iint E_{\sqrt{H}}(\lambda) \frac{d}{d\lambda}\left(\phi\left(\frac{\lambda}{2^m}\right) Q_j(\lambda)^*\right) \frac{d}{d\mu}\left(Q_j(\mu)\phi\left(\frac{\mu}{2^n}\right)\right) E_{\sqrt{H}}(\mu)\, d\lambda\, d\mu.$$

It is clear that each of these operators is uniformly bounded in m, n in operator norm. To apply Cotlar–Stein, we show a estimate of the form $C2^{-|m-n|}$ for the operator norm of this term. Write $Q^*_{j,m}(\lambda)$ and $Q_{j,n}(\mu)$ for the operators in parentheses above. Consider first the case $2 \leq j \leq N_l$, in which Q_j has Schwartz kernel supported near the boundary of the diagonal. For convenience of exposition, we assume that $\lambda, \mu \leq 2$ (or, equivalently, $m, n \leq 1$). Then, by the construction of Q_j for $2 \leq j \leq N_l$ (see Section 2D and (4-1)), the scattering pseudodifferential operators $Q^*_{j,m}(\lambda)$ and $Q_{j,n}(\mu)$ are smooth and compactly supported in x'/λ and x'/μ, respectively, and are microlocally supported near the characteristic set. More precisely, we see the composition of the two scattering pseudodifferential operators for $j \geq 2$ takes the form

$$\begin{aligned} Q^*_{j,m}(\lambda)Q_{j,n}(\mu) = \int & e^{-i\lambda((y-y')\cdot\eta+(\sigma-1)\nu)/x'} e^{i\mu((y'-y'')\cdot\eta'+(\sigma'-1)\nu')/x'} \\ & \times q_{j,m}\left(\lambda, y', \frac{x'}{\lambda}, \eta, \nu\right) q_{j,n}\left(\mu, y', \frac{x'}{\mu}, \eta', \nu'\right) dx'\, dy'\, d\eta\, d\nu\, d\eta'\, d\nu', \end{aligned}$$

where $\sigma = x'/x$ and $\sigma' = x'/x''$, and $q_{j,m}$ and $q_{j,n}$ are smooth and polyhomogeneous in λ and μ and compactly supported in x'/λ, x'/μ and y'. In addition, we have $\nu^2 + |\eta|^2 \geq \frac{1}{4}$ and $\nu'^2 + |\eta'|^2 \geq \frac{1}{4}$ on the support of $q_{j,m}q_{j,n}$. By symmetry, we assume $\lambda > \mu$ without loss of generality. Let us introduce the operator

$$\mathscr{L} = i[\lambda(|\nu|^2 + |\eta|^2)]^{-1}(x'\eta\,\partial_{y'} - \nu x'^2\,\partial_{x'});$$

then $\mathscr{L}e^{-i\lambda((y-y')\cdot\eta+(\sigma-1)\nu)/x'} = e^{-i\lambda((y-y')\cdot\eta+(\sigma-1)\nu)/x'}$. By using $\mathscr{L}$ to integrate by parts, we gain the factor λ^{-1} since $|\nu|^2 + |\eta|^2$ is uniformly bounded from below; we incur a factor μ if the derivative falls on $e^{i\mu((y'-y'')\cdot\eta'+(\sigma'-1)\nu')/x'}$, or a factor of x' or x'^2/μ if the derivative falls on $q_{j,m}$ or $q_{j,n}$. Since $x' \le \mu$ on the support of $q_{j,m}$, we have an overall gain of $\mu/\lambda \sim 2^{-|m-n|}$. The L^2 boundedness of the spectral projection gives $\|U^*_{j,m}(0)U_{j,n}(0)\|_{L^2\to L^2} \le C2^{-|m-n|}$.

A similar argument works if one or both of m and n are at least 1.

A similar estimate is true in the case $N_l + 1 \le j \le N$, in which case we are automatically in the high-energy case, and with Schwartz kernels supported in the interior of $M^\circ \times M^\circ$. The argument is also almost exactly the same as the previous case. We can write the composition

$$\frac{d}{d\lambda}\left(\phi\left(\frac{\lambda}{2^j}\right)Q_j(\lambda)^*\right)\frac{d}{d\mu}\left(Q_j(\mu)\phi\left(\frac{\mu}{2^k}\right)\right)$$

in the form

$$\lambda^n\mu^n \iiint e^{i\lambda(z-z'')\cdot\zeta} q_{j,m}(z'',\zeta,\lambda)e^{i\mu(z''-z')\cdot\zeta'} q_{j,n}(z'',\zeta',\mu)\,d\zeta\,d\zeta'\,dz'', \tag{5-9}$$

where $q_{j,m}$ is supported where $\lambda \sim 2^m$ and $|\zeta|^2 \sim 1$, and is such that $D^\alpha_z D^\beta_\zeta q_{j,m}$ is bounded by $C\lambda^{-1}$. Assume without loss of generality that $m > n$, i.e., $\lambda > \mu$ on the support of the integrand. We note that the differential operator

$$\mathscr{L} = \frac{i\zeta\cdot\partial_{z''}}{\lambda|\zeta|^2}$$

leaves $e^{i\lambda(z-z'')\cdot\zeta}$ invariant, so we can apply it to this phase factor in the integral (5-9). Integrating by parts, the $\partial_{z''}$ derivative either hits the other phase factor $e^{i\mu(z''-z')\cdot\zeta'}$, in which case we incur a factor of μ, or it hits one of the symbols $q_{i,j}$ or $q_{i,k}$, in which case we incur no factor. Thus, we gain a factor of either $\mu/\lambda \sim 2^{-|j-k|}$ or $1/\lambda$ — which is even better since $\mu > 1$ on the support of $q_{j,n}(z'',\zeta',\mu)$. This completes the Cotlar–Stein estimates for $U_i(0)$.

It now follows from the Cotlar–Stein lemma that $U_j(0)^*$, $j = 2, \dots, N$, is well-defined as the strong limit of the sequence of operators

$$\sum_{|m|\le l} U_{j,m}(0)^*.$$

Consider the sequence $\sum_{|m|\le l} U_{j,m}(t)^*$. We claim that this sequence converges strongly and define $U_j(t)^*$ to be this limit. To prove this claim, choose an arbitrary $f \in L^2(M^\circ)$. We have shown that

$$\lim_{l\to\infty}\sup_{L>l}\left\|\sum_{l\le|m|\le L} U_{j,m}(0)^* f\right\|_2^2 = 0.$$

This is equivalent to

$$\lim_{l\to\infty}\sup_{L>l}\sum_{l\le|m|,|m'|\le L}\langle U_{j,m}(0)U_{j,m'}(0)^* f, f\rangle = 0.$$

But we saw in (5-8) that $U_{j,m}(0)U_{j,m'}(0)^* = U_{j,m}(t)U_{j,m'}(t)^*$. Hence we have

$$\lim_{l\to\infty} \sup_{L>l} \sum_{l\le|m|,|m'|\le L} \langle U_{j,m}(t)U_{j,m'}(t)^* f, f\rangle = 0,$$

which implies that

$$\lim_{l\to\infty} \sup_{L>l} \Big\| \sum_{l\le|m|\le L} U_{j,m}(t)^* f \Big\|_2^2 = 0.$$

Hence the sequence $\sum_{|m|\le l} U_{j,m}(t)^* f$ converges for every $f \in L^2(M^\circ)$ as $l \to \infty$, i.e., the sequence $\sum_{|m|\le l} U_{j,m}(t)^*$ converges strongly. We see from this that the integral

$$\int e^{-it\lambda^2}\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j(\lambda)^*$$

converges in the strong topology, hence defines $U_j(t)^*$. Finally we show that the operator norm of $U_j(t)^*$ is bounded uniformly in t. Since $\sum_{|m|\le l} U_{j,m}(t)^*$ converges in the strong operator topology, we have

$$\|U_j(t)^*\| \le \sup_{l\to\infty} \Big\| \sum_{|m|\le l} U_{j,m}(t)^* \Big\|.$$

But we have

$$\Big\| \sum_{|m|\le l} U_{j,m}(t)^* \Big\|^2 = \Big\| \sum_{|m|,|m'|\le l} U_{j,m}(t)U_{j,m'}(t)^* \Big\| = \Big\| \sum_{|m|,|m'|\le l} U_{j,m}(0)U_{j,m'}(0)^* \Big\| = \Big\| \sum_{|m|\le l} U_{j,m}(0)^* \Big\|^2$$

and the operator norm of $\sum_{|m|\le l} U_{j,m}(0)^*$ is bounded uniformly in l by the estimates proved above using the Cotlar–Stein lemma.

This completes the proof of Proposition 5.1. □

Remark 5.4. This argument allows us to avoid using a Littlewood–Paley-type decomposition in this setting. Littlewood–Paley-type estimates were established in [Bouclet 2010] for asymptotically conic manifolds in the form of

$$\|f\|_{L^p} \lesssim \Big(\sum_{k\ge 0} \|\phi(2^{-2k}\Delta_g) f\|_{L^p}^2 \Big)^{\frac{1}{2}} + \Big\| \sum_{k\le 0} \phi(2^{-2k}\Delta_g) f \Big\|_{L^p}.$$

6. Dispersive estimates

In this section, we use stationary phase and Proposition 1.5 to establish the microlocalized dispersive estimates.

Proposition 6.1 (microlocalized dispersive estimates). *Let $Q_j(\lambda)$ be as defined in* (4-1). *Then, for all integers $j \ge 1$, the kernel estimate*

$$\Big| \int_0^\infty e^{it\lambda^2} (Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j^*(\lambda))(z, z')\, d\lambda \Big| \le C|t|^{-n/2} \tag{6-1}$$

holds for a constant C independent of points $z, z' \in M^\circ$.

Proof. The key to the proof is to use the estimates in Proposition 1.5. We first consider $j = 1$. Since the term with $Q_1(\lambda)$ satisfies (1-14) with only the "b" term, then we can use the estimate (1-16) to obtain

$$\left|\left(\frac{d}{d\lambda}\right)^N (Q_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_1^*(\lambda))(z,z')\right| \le C_N \lambda^{n-1-N} \quad \text{for all } N \in \mathbb{N}. \tag{6-2}$$

Let δ be a small constant to be chosen later. Recall that we chose $\phi \in C_c^\infty\big(\big[\frac{1}{2}, 2\big]\big)$ with $\sum_{m\in\mathbb{Z}} \phi(2^{-m}\lambda) = 1$; we write $\phi_0(\lambda) = \sum_{m\le -1} \phi(2^{-m}\lambda)$. Then

$$\left|\int_0^\infty e^{it\lambda^2} (Q_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_1^*(\lambda))(z,z') \phi_0\left(\frac{\lambda}{\delta}\right) d\lambda\right| \le C \int_0^\delta \lambda^{n-1}\, d\lambda \le C\delta^n.$$

We use integration by parts N times to obtain, using (6-2),

$$\begin{aligned}
&\left|\int_0^\infty e^{it\lambda^2} \sum_{m\ge 0} \phi\left(\frac{\lambda}{2^m\delta}\right) (Q_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_1^*(\lambda))(z,z')\, d\lambda\right| \\
&\qquad \le \sum_{m\ge 0} \left|\int_0^\infty \left(\frac{1}{2\lambda t}\frac{\partial}{\partial\lambda}\right)^N (e^{it\lambda^2}) \phi\left(\frac{\lambda}{2^m\delta}\right) (Q_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_1^*(\lambda))(z,z')\, d\lambda\right| \\
&\qquad \le C_N |t|^{-N} \sum_{m\ge 0} \int_{2^{m-1}\delta}^{2^{m+1}\delta} \lambda^{n-1-2N}\, d\lambda \\
&\qquad \le C_N |t|^{-N} \delta^{n-2N}.
\end{aligned}$$

Choosing $\delta = |t|^{-1/2}$, we have thus proved

$$\left|\int_0^\infty e^{it\lambda^2} (Q_1(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_1^*(\lambda))(z,z')\, d\lambda\right| \le C_N |t|^{-n/2}. \tag{6-3}$$

Now we consider the case $j \ge 2$. Let $r = d(z,z')$ and $\bar{r} = rt^{-1/2}$. In this case, we write the kernel using Proposition 1.5 as

$$\begin{aligned}
&\int_0^\infty e^{it\lambda^2} \big(Q_j(\lambda)\, dE_{\sqrt{\boldsymbol{H}}}(\lambda)\, Q_j^*(\lambda)\big)(z,z')\, d\lambda \\
&\qquad = \sum_\pm \int_0^\infty e^{it\lambda^2} e^{\pm ir\lambda} \lambda^{n-1} a_\pm(\lambda,z,z')\, d\lambda + \int_0^\infty e^{it\lambda^2} \lambda^{n-1} b(\lambda,z,z')\, d\lambda \\
&\qquad = t^{-n/2} \sum_\pm \int_0^\infty e^{i\lambda^2} e^{\pm i\bar{r}\lambda} \lambda^{n-1} a_\pm(t^{-1/2}\lambda,z,z')\, d\lambda + \int_0^\infty e^{it\lambda^2} \lambda^{n-1} b(\lambda,z,z')\, d\lambda,
\end{aligned} \tag{6-4}$$

where $a_\pm$ satisfies

$$|\partial_\lambda^\alpha a_\pm(\lambda,z,z')| \le C_\alpha \lambda^{-\alpha} (1 + \lambda d(z,z'))^{-(n-1)/2},$$

and therefore

$$|\partial_\lambda^\alpha (a_\pm(t^{-1/2}\lambda,z,z'))| \le C_\alpha \lambda^{-\alpha} (1 + \lambda\bar{r})^{-(n-1)/2}. \tag{6-5}$$

By (1.16), the above term with $b(\lambda,z,z')$ can be estimated by using the same argument as for Q_1. Now we consider first term in the right-hand side of (6-4). We divide it into two pieces using the partition of unity above. It suffices to prove that there exists a constant C independent of $\bar{r}$ such that

$$I^{\pm} := \left| \int_0^\infty e^{i\lambda^2} e^{\pm i\bar{r}\lambda} \lambda^{n-1} a_{\pm}(t^{-1/2}\lambda, z, z')\phi_0(\lambda)\, d\lambda \right| \leq C,$$
$$II^{\pm} := \left| \sum_{m\geq 0} \int_0^\infty e^{i\lambda^2} e^{\pm i\bar{r}\lambda} \lambda^{n-1} a_{\pm}(t^{-1/2}\lambda, z, z')\phi\left(\frac{\lambda}{2^m}\right) d\lambda \right| \leq C.$$

The estimate for $I^{\pm}$ is obvious, since $\lambda \leq 1$. For II^{+}, we use integration by parts. Notice that

$$L^{+}(e^{i\lambda^2 + i\bar{r}\lambda}) = e^{i\lambda^2 + i\bar{r}\lambda}, \quad L^{+} = \frac{-i}{2\lambda + \bar{r}} \frac{\partial}{\partial\lambda}.$$

Writing

$$e^{i\lambda^2 + i\bar{r}\lambda} = (L^{+})^N (e^{i\lambda^2 + i\bar{r}\lambda})$$

and integrating by parts, we gain a factor of λ^{-2N} thanks to (6-5). Thus II^{+} can be estimated by

$$\sum_{m\geq 0} \int_{\lambda\sim 2^m} \lambda^{n-1-2N}\, d\lambda \leq C.$$

To treat II^{-}, we introduce a further decomposition, based on the size of $\bar{r}\lambda$. We write $II^{-} = II_1^{-} + II_2^{-}$, where (dropping the $-$ superscripts and subscripts from here on)

$$II_1 = \left| \sum_{m\geq 0} \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')\phi\left(\frac{\lambda}{2^m}\right)\phi_0(4\bar{r}\lambda)\, d\lambda \right|,$$
$$II_2 = \left| \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')(1-\phi_0(\lambda))(1-\phi_0(4\bar{r}\lambda))\, d\lambda \right|.$$

Let $\Phi(\lambda, \bar{r}) = \lambda^2 - \bar{r}\lambda$. We first consider II_1. Since the integral for II_1 is supported where $\lambda \leq (4\bar{r})^{-1}$ and $\lambda \geq \frac{1}{2}$, the integrand is only nonzero when $\bar{r} \leq \frac{1}{2}$. Therefore, $|\partial_\lambda \Phi| = 2\lambda - \bar{r} \geq \frac{1}{2}\lambda$. Define the operator $L = L(\lambda, \bar{r}) = (2\lambda - \bar{r})^{-1}\, \partial_\lambda$. By (6-5) and using integration by parts, we obtain, for $N > \frac{1}{2}n$,

$$\begin{aligned} II_1 &\leq \sum_{m\geq 0} \left| \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')\phi\left(\frac{\lambda}{2^m}\right)\phi_0(4\bar{r}\lambda)\, d\lambda \right| \\ &= \sum_{m\geq 0} \left| \int L^N(e^{i(\lambda^2 - \bar{r}\lambda)}) \left[\lambda^{n-1} a(t^{-1/2}\lambda, z, z')\phi\left(\frac{\lambda}{2^m}\right)\phi_0(4\bar{r}\lambda)\right] d\lambda \right| \\ &\leq C_N \sum_{m\geq 0} \int_{|\lambda|\sim 2^m} \lambda^{n-1-2N}\, d\lambda \\ &\leq C_N. \end{aligned}$$

Finally, we consider II_2. Here, we replace the decomposition $\sum_m \phi(2^{-m}\lambda)$ with a different decomposition, based on the size of $\partial_\lambda \Phi$:

$$\begin{aligned} II_2 \le & \left| \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')(1-\phi_0(\lambda))\phi_0(2\lambda-\bar{r})(1-\phi_0(4\bar{r}\lambda))\, d\lambda \right| \\ & + \sum_{m\ge 0} \left| \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')(1-\phi_0(\lambda))\phi\left(\frac{2\lambda-\bar{r}}{2^m}\right)(1-\phi_0(4\bar{r}\lambda))\, d\lambda \right| \\ := & \, II_2^1 + II_2^2. \end{aligned}$$

If $\bar{r} \le 10$, then for the integrand of II_2^1 to be nonzero we must have $\lambda \le 10$, due to the ϕ_0 factor. Then it is easy to see that II_2^1 is uniformly bounded. If $\bar{r} \ge 10$, we have $\bar{r} \sim \lambda$ since $|2\lambda - \bar{r}| \le 1$. Hence, using (6-5) with $\alpha = 0$,

$$II_2^1 \le \int_{|2\lambda-\bar{r}|\le 1} \lambda^{n-1}(1+\bar{r}\lambda)^{-(n-1)/2}\, d\lambda \le C.$$

Now we consider the second term. Integrating by parts, we show by (6-5) that

$$\begin{aligned} II_2^2 & \le \sum_{m\ge 0} \left| \int e^{i\lambda^2} e^{-i\bar{r}\lambda} \lambda^{n-1} a(t^{-1/2}\lambda, z, z')(1-\phi_0(\lambda))\phi\left(\frac{2\lambda-\bar{r}}{2^m}\right)(1-\phi_0(4\bar{r}\lambda))\, d\lambda \right| \\ & = \sum_{m\ge 0} \left| \int L^N(e^{i(\lambda^2-\bar{r}\lambda)}) \left[\lambda^{n-1} a(t^{-1/2}\lambda, z, z')(1-\phi_0(\lambda))\phi\left(\frac{2\lambda-\bar{r}}{2^m}\right)(1-\phi_0(4\bar{r}\lambda)) \right] d\lambda \right| \\ & \le C_N \sum_{m\ge 0} 2^{-mN} \int_{|2\lambda-\bar{r}|\sim 2^m} \lambda^{n-1}(1+\bar{r}\lambda)^{-(n-1)/2}\, d\lambda. \end{aligned}$$

If $\bar{r} \le 2^{m+1}$, then $\lambda \le 2^{m+2}$ on the support of the integrand. The m-th term can then be estimated by $C_N 2^{-mN} 2^{(m+2)n}$, which is summable for $N > n$. Otherwise, we have $\lambda \sim \bar{r}$, which means the integrand is bounded and we estimate the m-th term by $C_N 2^{-mN} 2^m$, which is summable for $N > 1$. Therefore, we have completed the proof of Proposition 6.1. □

7. Homogeneous Strichartz estimates

We use the L^2 estimates and the microlocalized dispersive estimates to conclude the proof of Theorem 1.1. By Proposition 5.1, we have, for all $t \in \mathbb{R}$ and all $u_0 \in L^2$,

$$\|U_j(t)u_0\|_{L^2(M^\circ)} \lesssim \|u_0\|_{L^2(M^\circ)}.$$

By Lemma 5.3,

$$U_j(s)U_j^*(t)f = \int_0^\infty e^{i(s-t)\lambda^2} Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_j^*(\lambda) f.$$

Hence we have the decay estimates, by Proposition 6.1,

$$\|U_j(s)U_j^*(t)f\|_{L^\infty} \lesssim |t-s|^{-n/2} \|f\|_{L^1}.$$

As a consequence of the Keel–Tao abstract Strichartz estimate [1998], we have

$$\|U_j(t)u_0\|_{L^q(\mathbb{R};L^r(M^\circ))} \lesssim \|u_0\|_{L^2(M^\circ)}, \tag{7-1}$$

where (q, r) is sharp $\frac{n}{2}$-admissible, that is, $q, r \geq 2$, $(q, r, n) \neq (2, \infty, 2)$ and $\frac{2}{q} + \frac{n}{r} = \frac{n}{2}$. By the definition of $U_j(t)$ based on the construction of Q_j, we see that

$$e^{it\boldsymbol{H}} = \sum_{j=1}^{N} U_j(t). \tag{7-2}$$

Combining (7-1) and (7-2) proves the long-time homogeneous Strichartz estimate.

8. Inhomogeneous Strichartz estimates

In this section, we prove Theorem 1.2, including at the endpoint $(q, r) = (\tilde{q}, \tilde{r}) = (2, 2n/(n-2))$ for $n \geq 3$. Let $\boldsymbol{U}(t) = e^{it\boldsymbol{H}} : L^2 \to L^2$. We have already proved that

$$\|\boldsymbol{U}(t)u_0\|_{L^q_t L^r_z} \lesssim \|u_0\|_{L^2}$$

holds for all (q, r) satisfying (1-2). By duality, the estimate is equivalent to

$$\left\| \int_{\mathbb{R}} \boldsymbol{U}(t)\boldsymbol{U}^*(s)F(s)\, ds \right\|_{L^q_t L^r_z} \lesssim \|F\|_{L^{\tilde{q}'}_t L^{\tilde{r}'}_z},$$

where both (q, r) and $(\tilde{q}, \tilde{r})$ satisfy (1-2). By the Christ–Kiselev lemma [2001], we obtain, for $q > \tilde{q}'$,

$$\left\| \int_{s<t} \boldsymbol{U}(t)\boldsymbol{U}^*(s)F(s)\, ds \right\|_{L^q_t L^r_z} \lesssim \|F\|_{L^{\tilde{q}'}_t L^{\tilde{r}'}_z}. \tag{8-1}$$

Notice that $\tilde{q}' \leq 2 \leq q$; therefore, we have proved all inhomogeneous Strichartz estimates except the endpoint $(q, r) = (\tilde{q}, \tilde{r}) = (2, 2n/(n-2))$. To treat the endpoint, we need to show the bilinear form estimate

$$|T(F, G)| \leq \|F\|_{L^2_t L^{r'}_z} \|G\|_{L^2_t L^{r'}_z}, \tag{8-2}$$

where $r = 2n/(n-2)$ and $T(F, G)$ is the bilinear form

$$T(F, G) = \iint_{s<t} \langle \boldsymbol{U}(t)\boldsymbol{U}^*(s)F(s), G(t)\rangle_{L^2}\, ds\, dt. \tag{8-3}$$

Theorem 1.2 follows from:

Proposition 8.1. *There exists a partition of the identity $Q_j(\lambda)$ on $L^2(M^\circ)$ such that, with $U_j(t)$ defined as in* (5-1), *there exists a constant C such that, for each pair (j, k), either*

$$\iint_{s<t} \langle U_j(t)U_k^*(s)F(s), G(t)\rangle_{L^2}\, ds\, dt \leq C\|F\|_{L^2_t L^{r'}_z} \|G\|_{L^2_t L^{r'}_z} \tag{8-4}$$

$$\text{or} \quad \iint_{s>t} \langle U_j(t)U_k^*(s)F(s), G(t)\rangle_{L^2}\, ds\, dt \leq C\|F\|_{L^2_t L^{r'}_z} \|G\|_{L^2_t L^{r'}_z}. \tag{8-5}$$

Proof of Theorem 1.2 assuming Proposition 8.1. We have proved that, for all $1 \le j \le N$,

$$\|U_j(t)u_0\|_{L^2_t L^r_z} \lesssim \|u_0\|_{L^2};$$

hence it follows by duality that, for all $1 \le j, k \le N$,

$$\iint_{\mathbb{R}^2} \langle U_j(t)U_k^*(s)F(s), G(t)\rangle_{L^2}\, ds\, dt \le C\|F\|_{L^2_t L^{r'}_z} \|G\|_{L^2_t L^{r'}_z}. \tag{8-6}$$

Subtracting (8-5) from (8-6) shows that (8-4) holds for every pair (j, k). Then, by summing over all j and k, we obtain (8-2). □

To prove Proposition 8.1 we use the following lemma, proved in [Guillarmou and Hassell 2014, Lemmas 5.3 and 5.4].

Lemma 8.2. *The partition of the identity $Q_j(\lambda)$ can be chosen so that the pairs of indices (j, k), $1 \le j, k \le N$, can be divided into three classes,*

$$\{1, \dots, N\}^2 = J_{\text{near}} \cup J_{\text{not-out}} \cup J_{\text{not-inc}},$$

such that

- *if $(j, k) \in J_{\text{near}}$, then $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k(\lambda)^*$ satisfies the conclusions of Proposition 1.5*;
- *if $(j, k) \in J_{\text{non-inc}}$, then $Q_j(\lambda)$ is not incoming-related to $Q_k(\lambda)$, in the sense that no point in the operator wavefront set* (*microlocal support*) *of $Q_j(\lambda)$ is related to a point in the operator wavefront set of $Q_k(\lambda)$ by backward bicharacteristic flow*;
- *if $(j, k) \in J_{\text{non-out}}$, then $Q_j(\lambda)$ is not outgoing-related to $Q_k(\lambda)$, in the sense that no point in the operator wavefront set of $Q_j(\lambda)$ is related to a point in the operator wavefront set of $Q_k(\lambda)$ by forward bicharacteristic flow.*

We exploit the not-incoming or not-outgoing property of $Q_j(\lambda)$ with respect to $Q_k(\lambda)$ in the following two lemmas.

Lemma 8.3. *Let $Q_j(\lambda)$ and $Q_k(\lambda)$ be such that Q_j is not outgoing-related to Q_k. Then, for $\lambda \le 2$, as a multiple of $|dg\, dg'|^{1/2}\, |d\lambda|$ the Schwartz kernel of $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k(\lambda)^*$ can be expressed as the sum of a finite number of terms of the form*

$$\lambda^{n-1} \int_{\mathbb{R}^k} e^{i\lambda\Phi(y,y',\sigma,v)/x} \left(\frac{x'}{\lambda}\right)^{\frac{n-1}{2}-\frac{k}{2}} a\left(\lambda, y, y', \sigma, \frac{x'}{\lambda}, v\right) dv \tag{8-7}$$

$$\text{or} \quad \lambda^{n-1} \int_{\mathbb{R}^{k-1}} \int_0^\infty e^{i\lambda\Phi(y,y',\sigma,v,s)/x} \left(\frac{x'}{\lambda s}\right)^{\frac{n-1}{2}-\frac{k}{2}} s^{n-2} a\left(\lambda, y, y', \sigma, \frac{x'}{\lambda}, v, s\right) ds\, dv \tag{8-8}$$

in the region $\sigma = x/x' \le 2$, $x'/\lambda \le 2$, or

$$\lambda^{n-1} a\left(\lambda, y, y', \sigma, \frac{x'}{\lambda}\right) \tag{8-9}$$

in the region $\sigma = x/x' \le 2$, $x'/\lambda \ge 1$, where, in each case, $\Phi < -\epsilon < 0$ and a is a smooth function, compactly supported in the v and s variables (*where present*), *such that $|(\lambda\, \partial_\lambda)^N a| \le C_N$ for all $N \in \mathbb{N}$.*

In each case, we may assume that $k \le n-1$; *if* $k = 0$ *in* (8-7) *or* $k = 1$ *in* (8-8) *then there is no variable* v *and no* v *integral. The key point is that, in each expression, the phase function is strictly negative.*

If, instead, Q_j *is not incoming-related to* Q_k, *then the same conclusion holds with the reversed sign: the Schwartz kernel can be written as a finite sum of terms with a strictly positive phase function.*

Remark 8.4. For $\sigma \ge \frac{1}{2}$, the Schwartz kernel has a similar description, as follows immediately from the symmetry of the kernel under interchanging the left and right variables.

Proof. The statement that the Schwartz kernel has the indicated forms above follows immediately from the description of the spectral measure in [Guillarmou et al. 2013a, Theorem 3.10] as a Legendre distribution in the class $I^{m,p;r_{\mathrm{lb}},r_{\mathrm{rb}}}(M^2_{k,b}, (L^{\mathrm{bf}}, L^\sharp); \Omega^{1/2}_{k,b})$, where $m = -\frac{1}{2}$, $p = \frac{1}{2}(n-2)$ and $r_{\mathrm{lb}} = r_{\mathrm{rb}} = \frac{1}{2}(n-1)$. The bound on k follows from the fact that k can be taken as the drop in rank of the projection from L^{bf} to the base $(\partial M)^2 \times (0,\infty)_\sigma$, which is the front face (that is, the face created by blow-up) of M^2_{b}. We claim that the drop in rank is at most $n-1$, which proves that we may assume that $k \le n-1$. To prove this claim, we show that the differentials $dy_1, \dots dy_{n-1}$ and at least one of $d\sigma, dy'_1, \dots, dy'_{n-1}$ are linearly independent on L. This can be seen from the description of L as the flowout from the set

$$\{(y, y, 1, \mu, -\mu, \nu, -\mu) \mid \nu^2 + h(\mu) = 1\}, \tag{8-10}$$

using the coordinates of (2-6), by the flow of the vector field V_r, which is the vector field given by x^{-1} multiplied by the Hamilton vector field of the principal symbol of Δ acting in the right variables on $M^2_{k,b}$. In fact, $V_r = \sin s' \, \partial_{s'}$ in the coordinates (s, s') on the leaves γ^2 of (2-6) and takes the form (see [Hassell and Vasy 2001, Equation (2.26)] or [Guillarmou et al. 2013a, Equation (3.5)])

$$2\nu'\sigma \frac{\partial}{\partial\sigma} - 2\nu'\mu' \cdot \frac{\partial}{\partial\mu'} + h' \frac{\partial}{\partial\nu'} + \left(\frac{\partial h'}{\partial\mu'} \frac{\partial}{\partial y'} - \frac{\partial h'}{\partial y'} \frac{\partial}{\partial\mu'} \right), \quad h' = h(y', \mu') = \sum_{i,j} h^{ij}(y') \mu'_i \mu'_j.$$

It is clear that $dy_1, \dots, dy_{n-1}$ are linearly independent at the initial set (8-10). Moreover, their Lie derivative with respect to V_r vanishes, so they are linearly dependent on all of L^{bf}. Also, since $h' + \nu'^2 = 1$ on L^{bf}, either the ∂_σ or the $\partial_{y'}$ component of the vector field V_r does not vanish, unless $\sigma = 0$, showing that either $d\sigma$ or one of the dy'_i do not vanish at each point of L^{bf} for $\sigma \ne 0$. But it was shown in [Hassell and Vasy 2001] that L^{bf} is transversal to the boundary at $\sigma = 0$, which means that $d\sigma \ne 0$ on L^{bf} when σ is small. This proves the claim.

We next show that Φ can be taken to be strictly negative. We use the microlocal support estimates from [Guillarmou et al. 2013b]. Applying [Guillarmou et al. 2013b, Corollary 5.3], we find that the microlocal support of $Q_j(\lambda) \, dE_{\sqrt{H}}(\lambda) \, Q_k(\lambda)^*$ is contained in that part of L^{bf} where, in the notation of (2-6), $s < s'$ (since the initial set (8-10) corresponds to $s = s'$, and ∂_s and $\partial_{s'}$ move in the outgoing and incoming directions, respectively, along the flow). Repeating the calculation following (2-6), we see that the value of Φ "on the Legendrian" is $\Phi = -\cos s + \sigma \cos s' = (\sin s')^{-1} \sin(s - s')$, which is strictly negative. By restricting the support of the amplitude a in (8-7)–(8-9), we can assume that Φ is negative everywhere on the support of the integrand. □

Lemma 8.5. *Let $Q_j(\lambda)$ and $Q_k(\lambda)$ be such that Q_j is not outgoing-related to Q_k. Then, for $\lambda \geq 1$, and as a multiple of $|dg\, dg'|^{1/2}\,|d\lambda|$, the Schwartz kernel of $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k(\lambda)^*$ can be written in terms of a finite number of oscillatory integrals of the form*

$$\int_{\mathbb{R}^k} e^{i\lambda\Phi(y,y',\sigma,x,v)/x}\lambda^{n-1+k/2}x^{(n-1)/2-k/2}a(\lambda, y, y', \sigma, x, v)\, dv \tag{8-11}$$

$$\text{or}\quad \int_{\mathbb{R}^{k-1}}\int_0^\infty e^{i\lambda\Phi(y,y',\sigma,x,v,s)/x}\lambda^{n-1+k/2}\Big(\frac{x}{s}\Big)^{(n-1)/2-k/2}s^{n-2}a(\lambda, y, y', \sigma, x, v, s)\, ds\, dv \tag{8-12}$$

in the region $\sigma = x/x' \leq 2$, $x \leq \delta$, or

$$\int_{\mathbb{R}^k} e^{i\lambda\Phi(z,z',v)}\lambda^{n-1+k/2}a(\lambda, z, z', v)\, dv \tag{8-13}$$

in the region $x \geq \delta$, $x' \geq \delta$, where, in each case, $\Phi < -\epsilon < 0$ and a is a smooth function compactly supported in the v and s variables (where present) such that $|(\lambda\,\partial_\lambda)^N a| \leq C_N$. In each case, we may assume that $k \leq n-1$; if $k = 0$ in (8-11) *or* (8-13), *or $k = 1$ in* (8-12), *then there is no variable v and no v integral. Again, the key point is that, in each expression, the phase function is strictly negative.*

If, instead, Q_j is not incoming-related to Q_k, then the same conclusion holds with the reversed sign: the Schwartz kernel can be written as a finite sum of terms with a strictly positive phase function.

Proof. The proof is essentially identical to that of Lemma 8.3. The form of the oscillatory integrals comes from the fact that the spectral measure, for high energies, is a Legendre distribution in the class $I^{m,p;r_{\mathrm{lb}},r_{\mathrm{rb}}}(X, (L, L^\sharp); \Omega^{s\Phi}\Omega^{1/2})$, where the Lagrangian L is given by (3-3). The non-outgoing relation implies, via the microlocal support estimates of [Guillarmou et al. 2013b, Section 7], that $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k(\lambda)^*$ is microsupported where $\tau < 0$ in the coordinates of (3-3). Since $\Phi = \tau$ when $d_v\Phi = 0$, this implies that $\Phi < 0$ when $d_v\Phi = 0$. By restricting the support of the amplitude close to the set where $d_v\Phi = 0$, we can assume that $\Phi < 0$ everywhere on the support of the integrand. ☐

Lemma 8.6. *We have the following dispersive estimates on $U_j(t)U_k(s)^*$:*

- *If $(j, k) \in J_{\mathrm{near}}$, then for all $t \neq s$ we have*

$$\|U_j(t)U_k^*(s)\|_{L^1\to L^\infty} \leq C|t-s|^{-n/2}. \tag{8-14}$$

- *If (j, k) is such that Q_j is not outgoing-related to Q_k, and $t < s$, then*

$$\|U_j(t)U_k^*(s)\|_{L^1\to L^\infty} \leq C|t-s|^{-n/2}. \tag{8-15}$$

- *Similarly, if (j, k) is such that Q_j is not incoming-related to Q_k and $s < t$, then*

$$\|U_j(t)U_k^*(s)\|_{L^1\to L^\infty} \leq C|t-s|^{-n/2}. \tag{8-16}$$

Proof. The estimate (8-14) is essentially proved in Proposition 6.1, since we can use Proposition 1.5. Assume that Q_j is not incoming-related to Q_k and consider (8-16). By Lemma 5.3, $U_j(t)U_k(s)^*$ is given by

$$\int_0^\infty e^{i(t-s)\lambda^2}(Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k^*(\lambda))(z, z'). \tag{8-17}$$

Then we need to show that, for $s < t$,

$$\left| \int_0^\infty e^{i(t-s)\lambda^2} (Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k^*(\lambda))(z, z')\, d\lambda \right| \le C|t-s|^{-n/2}. \tag{8-18}$$

Case 1: $t - s \ge 1$**.** We introduce a dyadic partition of unity in λ. Let $\phi \in C_c^\infty\big(\big[\frac{1}{2}, 2\big]\big)$ be as in Section 5 with $\sum_m \phi(2^{-m}\sqrt{t-s}\lambda) = 1$, define

$$\phi_0(\sqrt{t-s}\lambda) = \sum_{m \le 0} \phi(2^{-m}\sqrt{t-s}\lambda)$$

and insert

$$1 = \phi_0(\sqrt{t-s}\lambda) + \sum_{m \ge 1} \phi_m(\sqrt{t-s}\lambda), \quad \phi_m(\lambda) := \phi(2^{-m}\lambda),$$

into the integral (8-17). In addition, we substitute for $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k^*(\lambda)$ one of the expressions in Lemmas 8.3 and 8.5. Since $t - s \ge 1$, for the ϕ_0 term only the low energy expressions are relevant. The estimate follows immediately from noticing that these expressions are pointwise bounded by $C\lambda^{n-1}$, using the fact that $k \le n-1$ in these expressions.

To treat the ϕ_m terms for $m \ge 1$, we again substitute one of the expressions in Lemmas 8.3 and 8.5. For notational simplicity we consider the expression (8-13), but the argument is similar in the other cases. We scale the λ variable and obtain the expression

$$\begin{aligned} &\int_0^\infty \int_{\mathbb{R}^k} e^{i(t-s)\lambda^2} e^{i\lambda\Phi(z,z',v)} \lambda^{n-1+k/2} a(\lambda, z, z', v) \phi_m(\sqrt{t-s}\lambda)\, dv\, d\lambda \\ &\quad = (t-s)^{-n/2-k/4} \int_0^\infty \int_{\mathbb{R}^k} e^{i(\bar\lambda^2 + \bar\lambda\Phi(z,z',v)/\sqrt{t-s})} \bar\lambda^{n-1+k/2} a\bigg(\frac{\bar\lambda}{\sqrt{t-s}}, y, y', \sigma, v\bigg) \phi_m(\bar\lambda)\, dv\, d\bar\lambda, \end{aligned} \tag{8-19}$$

where $\bar\lambda = \sqrt{t-s}\lambda$. We observe that the overall exponential factor is invariant under the differential operator

$$L = \frac{-i}{2\bar\lambda^2 + \bar\lambda\Phi/\sqrt{t-s}} \bar\lambda \frac{\partial}{\partial\bar\lambda}.$$

The adjoint of this is

$$L^t = -L + \frac{i}{2\bar\lambda^2 + \bar\lambda\Phi/\sqrt{t-s}} - i\frac{4\bar\lambda^2 + \bar\lambda\Phi/\sqrt{t-s}}{(2\bar\lambda^2 + \bar\lambda\Phi/\sqrt{t-s})^2}.$$

We apply L^N to the exponential factors and integrate by parts N times. Since $\Phi \ge 0$ according to Lemma 8.5, and since we have an estimate $|(\bar\lambda\, \partial_{\bar\lambda})^N a| \le C_N$, each time we integrate by parts we gain a factor $\bar\lambda^{-2} \sim 2^{-2m}$. It follows that the integral with $\phi(2^{-m}\bar\lambda)$ inserted is bounded by $(t-s)^{-n/2} 2^{-m(2N-n-k/2)}$ uniformly for $t - s \ge 1$. Hence we prove (8-16) by summing over $m \ge 0$. The argument to prove (8-15) is analogous.

Case 2: $t - s \le 1$**.** In this case, we use a dyadic decomposition in terms of the original variable λ. We consider the integral (8-17), insert the dyadic decomposition

$$1 = \sum_{m \ge 0} \phi_m(\lambda),$$

and substitute for $Q_j(\lambda)\, dE_{\sqrt{H}}(\lambda)\, Q_k^*(\lambda)$ one of the expressions in Lemmas 8.3 and 8.5.

For the case $m = 0$, the estimate follows immediately from the uniform boundedness of (8-7)–(8-9). For the cases $m \geq 1$, we use the expressions in Lemma 8.5 and observe that the overall exponential factor is invariant under the differential operator

$$L = \frac{-i}{2(t-s)\lambda^2 + \lambda\Phi} \lambda \frac{\partial}{\partial\lambda}.$$

The adjoint of this is

$$L^t = -L + \frac{i}{2(t-s)\lambda^2 + \lambda\Phi} - i\frac{4(t-s)\lambda^2 + \lambda\Phi}{(2(t-s)\lambda^2 + \lambda\Phi)^2}.$$

We apply L to the exponential factors N times and integrate by parts. Since $\Phi \geq \epsilon > 0$ according to Lemma 8.5, and since we have an estimate $|(\lambda\, \partial_\lambda)^N a| \leq C_N$, each time we integrate by parts we gain a factor $\lambda^{-1} \sim 2^{-m}$. It follows that the integral with $\phi(2^{-m}\lambda)$ inserted is bounded by $2^{-m(N-n-k/2)}$ uniformly for $t - s \leq 1$. Hence we prove (8-16) by summing over $m \geq 0$. The argument to prove (8-15) is analogous. ☐

Remark 8.7. Notice that, in the cases (8-15) and (8-16), there is a lot of "slack" in the estimates. This is because the sign of $t - s$ has the favourable sign relative to the sign of the phase function, so that the overall phase in integrals such as (8-19) are never stationary. Then integration by parts give us more decay than needed to prove the estimates. This is important because it overcomes the growth of the spectral measure as $\lambda \to \infty$ at conjugate points: at pairs of conjugate points we have $k > 0$ and we see from, say, (8-13) that the spectral measure will not obey the localized (near the diagonal) estimates of Proposition 1.5, by a factor $\lambda^{k/2}$. The geometric meaning of k is the drop in rank of the projection from L down to M_b^2, hence is positive precisely at pairs of conjugate points.

We now complete the proof of Theorem 1.2 by proving Proposition 8.1.

Proof of Proposition 8.1. We use a partition of the identity as in Lemma 8.2. In the case that $(j,k) \in J_{\text{near}}$, we have the dispersive estimate (8-14). This allows us to apply the argument of [Keel and Tao 1998, Sections 4–7] to obtain (8-4). In the case that $(j,k) \in J_{\text{non-out}}$, we obtain (8-4) following the argument in [Keel and Tao 1998] since we have the dispersive estimate (8-16) when $s < t$. Finally, in the case that $(j,k) \in J_{\text{non-inc}}$, we obtain (8-5) since we have the dispersive estimate (8-15) for $s > t$. ☐

Remark 8.8. The endpoint inhomogeneous Strichartz estimate is closely related to the uniform Sobolev estimate

$$\|(H - \alpha)^{-1}\|_{L^r \to L^{r'}} \leq C, \quad r = \frac{2n}{n+2}, \tag{8-20}$$

where C is independent of $\alpha \in \mathbb{C}$. This estimate was proved by [Kenig et al. 1987] for the flat Laplacian and by [Guillarmou and Hassell 2014] for the Laplacian on nontrapping asymptotically conic manifolds (it was also shown in [Guillarmou and Hassell 2014] that (8-20) holds for $r \in [2n/(n+2), 2(n+1)/(n+3)]$ with a power of α on the right-hand side). In fact, it was pointed out to the authors by Thomas Duyckaerts and Colin Guillarmou that the endpoint inhomogeneous Strichartz estimate implies the uniform Sobolev estimate (8-20). To see this, we choose $w \in C_c^\infty(M^\circ)$ and $\chi(t)$ equal to 1 on $[-T, T]$ and zero for $|t| \geq T+1$ and let $u(t,z) = \chi(t)e^{i\alpha t}w(z)$. Then

$$(i\,\partial_t + \boldsymbol{H})u = F(t,z), \quad F(t,z) := \chi(t)e^{i\alpha t}(\boldsymbol{H}-\alpha)w(z) + i\chi'(t)e^{i\alpha t}w(z).$$

Applying the endpoint inhomogeneous Strichartz estimate, we obtain

$$\|u\|_{L_t^2 L_z^{r'}} \leq C\|F\|_{L_t^2 L_z^r}.$$

From the specific form of u and F we have

$$\|u\|_{L_t^2 L_z^{r'}} = \sqrt{2T}\|w\|_{L^{r'}} + O(1), \quad \|F\|_{L_t^2 L_z^r} = \sqrt{2T}\|(\boldsymbol{H}-\alpha)w\|_{L^r} + O(1).$$

Taking the limit $T \to \infty$ we find that

$$\|w\|_{L^{r'}} \leq C\|(\boldsymbol{H}-\alpha)w\|_{L^r},$$

which implies the uniform Sobolev estimate.

In the other direction, suppose that the uniform Sobolev estimate holds. If u and F satisfy (1-10), then taking the Fourier transform in t we find that

$$(\boldsymbol{H}-\alpha)\hat{u}(\alpha,z) = \widehat{F}(\alpha,z). \tag{8-21}$$

Suppose for a moment that the following statement were true: “Fourier transformation in t is a bounded linear map from $L^2(\mathbb{R}_t; L^p(M^\circ))$ to $L^2(\mathbb{R}_\alpha; L^p(M^\circ))$ for $p = r', r$”. Using this and the uniform Sobolev inequality, applied to (8-21), we would obtain the inhomogeneous Strichartz estimate. Unfortunately, the statement in quotation marks is known to be false, so this argument is purely heuristic. Nevertheless, it illustrates the close relation between the two estimates. It would be interesting to know if there are general conditions under which the two estimates are equivalent.

Acknowledgements

We thank Colin Guillarmou, Adam Sikora, Jean-Marc Bouclet, Thomas Duyckaerts and Pierre Portal for helpful conversations. This research was supported by Future Fellowship FT0990895 and Discovery Grants DP1095448 and DP120102019 from the Australian Research Council. Zhang was supported by Beijing Natural Science Foundation (1144014) and National Natural Science Foundation of China (11401024).

References

[Blair et al. 2008] M. D. Blair, H. F. Smith, and C. D. Sogge, “On Strichartz estimates for Schrödinger operators in compact manifolds with boundary”, *Proc. Amer. Math. Soc.* **136**:1 (2008), 247–256. MR 2008k:35386 Zbl 1169.35012

[Blair et al. 2009] M. D. Blair, H. F. Smith, and C. D. Sogge, “Strichartz estimates for the wave equation on manifolds with boundary”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:5 (2009), 1817–1829. MR 2010k:35500 Zbl 1198.58012

[Blair et al. 2012] M. D. Blair, G. A. Ford, S. Herr, and J. L. Marzuola, “Strichartz estimates for the Schrödinger equation on polygonal domains”, *J. Geom. Anal.* **22**:2 (2012), 339–351. MR 2891729 Zbl 1252.35237

[Bouclet 2010] J.-M. Bouclet, “Littlewood–Paley decompositions on manifolds with ends”, *Bull. Soc. Math. France* **138**:1 (2010), 1–37. MR 2011e:42020 Zbl 1198.42013

[Bouclet 2011] J.-M. Bouclet, “Strichartz estimates on asymptotically hyperbolic manifolds”, *Anal. PDE* **4**:1 (2011), 1–84. MR 2012i:58022 Zbl 1230.35027

[Bouclet and Tzvetkov 2007] J.-M. Bouclet and N. Tzvetkov, “Strichartz estimates for long range perturbations”, *Amer. J. Math.* **129**:6 (2007), 1565–1609. MR 2009d:58035

[Bouclet and Tzvetkov 2008] J.-M. Bouclet and N. Tzvetkov, “On global Strichartz estimates for non-trapping metrics”, *J. Funct. Anal.* **254**:6 (2008), 1661–1682. MR 2009d:35039 Zbl 1168.35005

[Bourgain 1999] J. Bourgain, *Global solutions of nonlinear Schrödinger equations*, American Mathematical Society Colloquium Publications **46**, American Mathematical Society, Providence, RI, 1999. MR 2000h:35147 Zbl 0933.35178

[Burq et al. 2004a] N. Burq, P. Gérard, and N. Tzvetkov, “Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds”, *Amer. J. Math.* **126**:3 (2004), 569–605. MR 2005h:58036 Zbl 1067.58027

[Burq et al. 2004b] N. Burq, F. Planchon, J. G. Stalker, and A. S. Tahvildar-Zadeh, “Strichartz estimates for the wave and Schrödinger equations with potentials of critical decay”, *Indiana Univ. Math. J.* **53**:6 (2004), 1665–1680. MR 2005k:35241 Zbl 1084.35014

[Burq et al. 2010] N. Burq, C. Guillarmou, and A. Hassell, “Strichartz estimates without loss on manifolds with hyperbolic trapped geodesics”, *Geom. Funct. Anal.* **20**:3 (2010), 627–656. MR 2012f:58068 Zbl 1206.58009

[Christ and Kiselev 2001] M. Christ and A. Kiselev, “Maximal functions associated to filtrations”, *J. Funct. Anal.* **179**:2 (2001), 409–425. MR 2001i:47054 Zbl 0974.47025

[Fanelli et al. 2013] L. Fanelli, V. Felli, M. A. Fontelos, and A. Primo, “Time decay of scaling critical electromagnetic Schrödinger flows”, *Comm. Math. Phys.* **324**:3 (2013), 1033–1067. MR 3123544 Zbl 1293.35266

[Ford 2010] G. A. Ford, “The fundamental solution and Strichartz estimates for the Schrödinger equation on flat Euclidean cones”, *Comm. Math. Phys.* **299**:2 (2010), 447–467. MR 2011i:35204 Zbl 1198.35062

[Ginibre and Velo 1985] J. Ginibre and G. Velo, “Scattering theory in the energy space for a class of nonlinear Schrödinger equations”, *J. Math. Pures Appl.* (9) **64**:4 (1985), 363–401. MR 87i:35171 Zbl 0535.35069

[Grafakos 2009] L. Grafakos, *Modern Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **250**, Springer, New York, 2009. MR 2011d:42001 Zbl 1158.42001

[Guillarmou and Hassell 2008] C. Guillarmou and A. Hassell, “Resolvent at low energy and Riesz transform for Schrödinger operators on asymptotically conic manifolds, I”, *Math. Ann.* **341**:4 (2008), 859–896. MR 2009j:58043 Zbl 1141.58017

[Guillarmou and Hassell 2014] C. Guillarmou and A. Hassell, “Uniform Sobolev estimates for non-trapping metrics”, *J. Inst. Math. Jussieu* **13**:3 (2014), 599–632. MR 3211800 Zbl 1321.58013

[Guillarmou et al. 2013a] C. Guillarmou, A. Hassell, and A. Sikora, “Resolvent at low energy, III: The spectral measure”, *Trans. Amer. Math. Soc.* **365**:11 (2013), 6103–6148. MR 3091277 Zbl 1303.58010

[Guillarmou et al. 2013b] C. Guillarmou, A. Hassell, and A. Sikora, “Restriction and spectral multiplier theorems on asymptotically conic manifolds”, *Anal. PDE* **6**:4 (2013), 893–950. MR 3092733 Zbl 1293.35187

[Hassell and Vasy 1999] A. Hassell and A. Vasy, “The spectral projections and the resolvent for scattering metrics”, *J. Anal. Math.* **79** (1999), 241–298. MR 2001d:58034 Zbl 0981.58025

[Hassell and Vasy 2001] A. Hassell and A. Vasy, “The resolvent for Laplace-type operators on asymptotically conic spaces”, *Ann. Inst. Fourier* (*Grenoble*) **51**:5 (2001), 1299–1346. MR 2002i:58037

[Hassell and Wunsch 2005] A. Hassell and J. Wunsch, “The Schrödinger propagator for scattering metrics”, *Ann. of Math.* (2) **162**:1 (2005), 487–523. MR 2006k:58048 Zbl 1126.58016

[Hassell and Wunsch 2008] A. Hassell and J. Wunsch, “The semiclassical resolvent and the propagator for non-trapping scattering metrics”, *Adv. Math.* **217**:2 (2008), 586–682. MR 2009b:58070 Zbl 1131.58018

[Hassell et al. 2005] A. Hassell, T. Tao, and J. Wunsch, “A Strichartz inequality for the Schrödinger equation on nontrapping asymptotically conic manifolds”, *Comm. Partial Differential Equations* **30**:1–3 (2005), 157–205. MR 2006i:58045 Zbl 1068.35119

[Hassell et al. 2006] A. Hassell, T. Tao, and J. Wunsch, "Sharp Strichartz estimates on nontrapping asymptotically conic manifolds", *Amer. J. Math.* **128**:4 (2006), 963–1024. MR 2007d:58053 Zbl 1177.58019

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudodifferential operators*, Grundlehren der Math. Wissenschaften **274**, Springer, Berlin, 1985. MR 87d:35002a Zbl 0601.35001

[Ivanovici 2010] O. Ivanovici, "On the Schrödinger equation outside strictly convex obstacles", *Anal. PDE* **3**:3 (2010), 261–293. MR 2011j:58037 Zbl 1222.35186

[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

[Kenig et al. 1987] C. E. Kenig, A. Ruiz, and C. D. Sogge, "Uniform Sobolev inequalities and unique continuation for second order constant coefficient differential operators", *Duke Math. J.* **55**:2 (1987), 329–347. MR 88d:35037 Zbl 0644.35012

[Marzuola et al. 2008] J. Marzuola, J. Metcalfe, and D. Tataru, "Strichartz estimates and local smoothing estimates for asymptotically flat Schrödinger equations", *J. Funct. Anal.* **255**:6 (2008), 1497–1553. MR 2011c:35485 Zbl 1180.35187

[Marzuola et al. 2010] J. Marzuola, J. Metcalfe, D. Tataru, and M. Tohaneanu, "Strichartz estimates on Schwarzschild black hole backgrounds", *Comm. Math. Phys.* **293**:1 (2010), 37–83. MR 2010m:58043 Zbl 1202.35327

[Melrose 1994] R. B. Melrose, "Spectral and scattering theory for the Laplacian on asymptotically Euclidian spaces", pp. 85–130 in *Spectral and scattering theory* (Sanda, 1992), edited by M. Ikawa, Lecture Notes in Pure and Appl. Math. **161**, Dekker, New York, 1994. MR 95k:58168

[Metcalfe and Tataru 2012] J. Metcalfe and D. Tataru, "Global parametrices and dispersive estimates for variable coefficient wave equations", *Math. Ann.* **353**:4 (2012), 1183–1237. MR 2944027 Zbl 1259.35006

[Mizutani 2012] H. Mizutani, "Strichartz estimates for Schrödinger equations on scattering manifolds", *Comm. Partial Differential Equations* **37**:2 (2012), 169–224. MR 2876829 Zbl 1244.35020

[Robbiano and Zuily 2005] L. Robbiano and C. Zuily, *Strichartz estimates for Schrödinger equations with variable coefficients*, Mém. Soc. Math. Fr. (N.S.) **101–102**, Société Mathématique de France, Paris, 2005. MR 2006i:35047 Zbl 1097.35002

[Sher 2013] D. A. Sher, "The heat kernel on an asymptotically conic manifold", *Anal. PDE* **6**:7 (2013), 1755–1791. MR 3148066 Zbl 1296.58015

[Staffilani and Tataru 2002] G. Staffilani and D. Tataru, "Strichartz estimates for a Schrödinger operator with nonsmooth coefficients", *Comm. Partial Differential Equations* **27**:7–8 (2002), 1337–1372. MR 2003f:35248 Zbl 1010.35015

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR 58 #23577 Zbl 0372.35001

[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

[Wunsch and Zworski 2000] J. Wunsch and M. Zworski, "Distribution of resonances for asymptotically Euclidean manifolds", *J. Differential Geom.* **55**:1 (2000), 43–82. MR 2002e:58062 Zbl 1030.58024

[Zhang 2015a] J. Zhang, "Linear restriction estimates for Schrödinger equation on metric cones", *Comm. Partial Differential Equations* **40**:6 (2015), 995–1028. MR 3321069 Zbl 1326.42029

[Zhang 2015b] J. Zhang, "Strichartz estimates and nonlinear wave equation on nontrapping asymptotically conic manifolds", *Adv. Math.* **271** (2015), 91–111. MR 3291858

Received 12 Apr 2015. Revised 24 Sep 2015. Accepted 28 Oct 2015.

ANDREW HASSELL: Andrew.Hassell@anu.edu.au
Department of Mathematics, Australian National University, Canberra ACT 2601, Australia

JUNYONG ZHANG: zhang_junyong@bit.edu.cn
Department of Mathematics, Beijing Institute of Technology, Beijing, 100081, China
and
Department of Mathematics, Australian National University, Canberra ACT 2601, Australia

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.193

LIMITING DISTRIBUTION OF ELLIPTIC HOMOGENIZATION ERROR WITH PERIODIC DIFFUSION AND RANDOM POTENTIAL

WENJIA JING

We study the limiting probability distribution of the homogenization error for second order elliptic equations in divergence form with highly oscillatory periodic conductivity coefficients and highly oscillatory stochastic potential. The effective conductivity coefficients are the same as those of the standard periodic homogenization, and the effective potential is given by the mean. We show that the limiting distribution of the random part of the homogenization error, as random elements in proper Hilbert spaces, is Gaussian and can be characterized by the homogenized Green's function, the homogenized solution and the statistics of the random potential. This generalizes previous results in the setting with slowly varying diffusion coefficients, and the current setting with fast oscillations in the differential operator requires new methods to prove compactness of the probability distributions of the random fluctuation.

1. Introduction

In this article we study the limiting distribution, in certain Hilbert spaces, of the homogenization error for second order elliptic equations in divergence form with highly oscillatory periodic diffusion coefficients and highly oscillatory random potential.

More precisely, we consider the following Dirichlet problem on an open bounded subset $D \subset \mathbb{R}^n$, with homogeneous boundary condition and a source term $f \in L^2(D)$,

$$\begin{cases} -\dfrac{\partial}{\partial x_i}\Big(a_{ij}\Big(\dfrac{x}{\varepsilon}\Big)\dfrac{\partial u^\varepsilon}{\partial x_j}(x,\omega)\Big) + q\Big(\dfrac{x}{\varepsilon},\omega\Big)u^\varepsilon(x,\omega) = f(x), & x \in D, \\ u^\varepsilon(x) = 0, & x \in \partial D. \end{cases} \tag{1-1}$$

The conductivity coefficients $\big(a_{ij}\big(\frac{\cdot}{\varepsilon}\big)\big)$ and the potential $q\big(\frac{\cdot}{\varepsilon},\omega\big)$ are highly oscillatory in space, and $0 < \varepsilon \ll 1$ indicates the small scale on which these coefficients oscillate. We assume that the conductivity coefficients are deterministic and periodic, and the potential is a stationary random field on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. More precise assumptions are given in Section 2. It is well known that, under mild assumptions like stationary ergodicity of $q(x,\omega)$, the equation above homogenizes; i.e., u^ε converges, almost surely in Ω, weakly in $H^1(D)$ and strongly in $L^2(D)$ to the solution of the deterministic homogenized problem

$$\begin{cases} -\bar{a}_{ij}\dfrac{\partial^2 u}{\partial x_i\,\partial x_j}(x) + \bar{q}u(x) = f(x), & x \in D, \\ u(x) = 0, & x \in \partial D. \end{cases} \tag{1-2}$$

MSC2010: primary 35R60; secondary 60B12.
Keywords: periodic and stochastic homogenization, random field, probability measures on Hilbert space, weak convergence of probability distributions.

Here, the effective conductivity coefficients $(\bar{a}_{ij})$ are constants defined by

$$\bar{a}_{ij} = \int_{\mathbb{T}^d} a_{ik}(y)\left(\delta_{kj} + \frac{\partial \chi^k}{\partial x_j}(y)\right) dy, \tag{1-3}$$

where $\mathbb{T}^d = [0,1]^d$ denotes the unit cell and the correctors χ^k, with $k = 1, \dots, d$, are given by the unique solution of the corrector equation

$$-\frac{\partial}{\partial x_i}\left(a_{ij}(y)\left(e_k + \frac{\partial \chi^k}{\partial x_j}(y)\right)\right) = 0 \quad \text{on } \mathbb{T}^d, \tag{1-4}$$

with the normalization condition $\int_{\mathbb{T}^d} \chi^k \, dy = 0$; e_k above is the k-th standard unit basis vector of $\mathbb{R}^d$. We note that this formula for $(\bar{a}_{ij})$ is exactly the classic periodic homogenization formula for effective conductivity. The effective potential $\bar{q}$ in (1-2) is given by the constant

$$\bar{q} = \mathbb{E} q(0, \omega), \tag{1-5}$$

where $\mathbb{E}$ denotes the mathematical mean with respect to $\mathbb{P}$.

In this paper we study the law (probability distribution) of the homogenization error $u^\varepsilon - u$, viewed as random elements in certain Hilbert spaces. We split this error into two parts: $\mathbb{E}u^\varepsilon - u$ and $u^\varepsilon - \mathbb{E}u^\varepsilon$. In view of the deterministic oscillations in the diffusion coefficients, we expect that the periodic homogenization error, in the replacement of $\left(a^{ij}\left(\frac{\cdot}{\varepsilon}\right)\right)$ to $(\bar{a}_{ij})$, makes significant contributions to the deterministic error $\mathbb{E}u^\varepsilon - u$. Indeed, we show later that this error is essentially of order $O(\varepsilon)$, the same as periodic homogenization. On the other hand, the effect of the random potential $q\left(\frac{\cdot}{\varepsilon}, \omega\right)$ becomes visible in the random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$, in which the (large) mean is removed. We are interested in characterizing the size and the law of this random fluctuation, and the answers depend on finer information of the random potential q, such as the decay rate of the correlations in q and higher-order moments of q; see Section 2 for notations and definitions.

We find that, when $q(x, \omega)$ has short-range correlations, the random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$ scales like $\varepsilon^{d/2 \wedge 2}$ in the $L^1(\Omega, L^2(D))$-norm, and scales like $\varepsilon^{d/2}$ when integrated against a test function. Moreover, the law of the scaled random fluctuation $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}u^\varepsilon)$ in $L^2(D)$ for $d = 2, 3$ and in $H^{-1}(D)$ for $d = 4, 5$ converges to Gaussian distributions as follows (see Theorem 2.4 for details):

$$\frac{u^\varepsilon - \mathbb{E}u^\varepsilon}{\sqrt{\varepsilon^d}} \xrightarrow{\text{distribution}} \sigma \int_D G(x, y) u(y) \, dW(y).$$

Here, $W(y)$ is the standard multiparameter Wiener process, and hence the law of the right-hand side above defines a Gaussian probability measure on $L^2(D)$ or $H^{-1}(D)$. This Gaussian distribution is determined by $G(x, y)$, which is the Green's function associated to the homogenized problem (1-2), $u(y)$, which is the homogenized solution, and σ, which is some statistical parameter of the random potential $q(x, \omega)$.

We also consider the case when $q(x, \omega) = \Phi(g(x, \omega))$ is constructed as a function of a Gaussian random field $g(x, \omega)$, and g has long-range correlations that decay like $|x|^{-\alpha}$, with $0 < \alpha < d$. Then the random fluctuation scales like $\varepsilon^{\alpha/2 \wedge 2}$ in the $L^1(\Omega, L^2(D))$-norm, and scales like $\varepsilon^{\alpha/2}$ when integrated against a test function. Moreover, the law of the scaled random fluctuation $\varepsilon^{-\alpha/2}(u^\varepsilon - \mathbb{E}u^\varepsilon)$ in $L^2(D)$ for $d = 2, 3$

and in $H^{-1}(D)$ for $d = 4, 5$ converges to a Gaussian distribution that can be written as a stochastic integral as above, but with dW replaced by $\dot{W}^\alpha\, dy$, where $\dot{W}^\alpha$ is a centered Gaussian random field with correlation function $|x-y|^{-\alpha}$, and σ is replaced by some other statistical parameter; see Theorem 6.2 for details.

The study of the limiting distribution of the homogenization error goes back to [Figari et al. 1982], where the Laplace operator with a random potential formed by Poisson bumps was considered. General random potential with short-range correlations was considered recently in [Bal 2008], and in [Bal and Jing 2010; 2011] for other nonoscillatory differential operators with random potential. Long-range correlated random potential was considered in [Bal et al. 2012]. When oscillatory differential operators were considered, the limiting distribution of homogenization error was obtained in [Bourgeat and Piatnitski 1999] for short-range correlated elliptic coefficients, and in [Bal et al. 2008] for the long-range correlated case, all in the one-dimensional setting. The main results of this paper show that the general framework developed in [Bal 2008; Bal and Jing 2011; Bal et al. 2012], in order to characterize the random fluctuation caused by the random potential, applies even when there are oscillations in the differential operators, as long as these oscillations are not statistically related to those of the random potential.

Our approach is as follows: we introduce an auxiliary problem with periodic diffusion coefficients and homogenized potential; let v^ε be the solution. Then the deterministic homogenization error $\mathbb{E}u^\varepsilon - u$ is essentially characterized by $v^\varepsilon - u$, which amounts to classical periodic homogenization theory. The random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$ is then the same as $(u^\varepsilon - v^\varepsilon) - \mathbb{E}(u^\varepsilon - v^\varepsilon)$, which can be represented as a truncated Neumann series. The first term X^ε in this series contributes to the limiting distribution. By Prohorov's theorem, we need to show that the probability measures of $\{X^\varepsilon\}$ are tight in the proper Hilbert space, and that their characteristic functions converge. The latter is essentially the convergence in distribution of the integration of X^ε against test functions; in view of the uniform-in-ε estimates of the Green's functions associated to the oscillatory diffusion, this step is the same as the earlier setting with nonoscillatory diffusion. The role of oscillations in the diffusion, however, becomes prominent in the step of proving tightness of the measures of $\{X^\varepsilon\}$. The simple and natural method used in [Bal et al. 2012] fails completely; see Section 7 for details. New ideas are needed: we obtain tightness of the measures of $\{X^\varepsilon\}$ in $L^2(D)$ by controlling the mean square of the H^s-norm of $\{X^\varepsilon\}$ for some $0 < s < \frac{1}{2}$; similarly, we get tightness in $H^{-1}(D)$ by controlling the mean square of the H^{-s}-norm with $\frac{1}{2} < s < 1$. The constraints on the spatial dimension d arise naturally in the proof of such controls.

Our analysis relies on uniform estimates of the Green's function associated to the periodic homogenization problem; we refer to [Avellaneda and Lin 1987; 1991] for the classical results, and to [Kenig et al. 2012; 2014] for recent development in this direction. We refer to [Armstrong and Smart 2014; Armstrong et al. 2015; Marahrens and Otto 2015; Gloria and Otto 2014] for recent results on uniform estimates of the Green's function for equations with highly oscillatory random diffusion coefficients in spatial dimension higher than one. We remark also that in the random setting, the limiting distribution of the corrector function and that of the full random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$, in negative Hölder space, were obtained in [Mourrat and Nolen 2015] and [Gu and Mourrat 2015] respectively, in the discrete setting; see also [Mourrat and Otto 2014]. Such results are apparently more challenging to obtain, and the proofs require delicate calculus in the (infinite-dimensional) probability space.

The rest of this paper is organized as follows: In Section 2 we make precise the main assumptions on the parameters of the homogenization problem, in particular on the properties of the random potential, and state the main results in the short-range correlation setting. Homogenization of (1-1) and some useful results on periodic homogenization theory are recalled in Section 3. Sections 4 and 5 are devoted to the proofs of the main results, where we characterize how the random fluctuation scales in the energy norm and in the weak topology, and determine the limiting distribution of the scaled fluctuation. We present new methods to prove the tightness of the probability measures of the random fluctuations. In Section 6, we state and prove the corresponding results in the long-range correlation setting. We make some comments and further discussions in Section 7 and prove some technical results, such as tightness criteria for probability measures, in the Appendix.

2. Assumptions, preliminaries and main results

2A. ***Assumptions on the coefficients.*** Throughout this paper, we assume that the domain D in (1-1) is an open bounded set of $\mathbb{R}^d$ with $C^{1,1}$-boundary. The coefficients $a_{ij}\big(\frac{x}{\varepsilon}\big)$ and $q\big(\frac{x}{\varepsilon},\omega\big)$ are the scaled versions of $a_{ij}(x)$ and $q(x,\omega)$. We make the following main assumptions on a_{ij} and q.

Periodic diffusion coefficients. For the functions (a_{ij}), we assume:

(A1) (periodicity) The function $A := (a_{ij}) : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is periodic. That is, for all $x \in \mathbb{R}^d$, $k \in \mathbb{Z}^d$ and $i, j = 1, 2, \dots, d$, we have

$$a_{ij}(x+k) = a_{ij}(x). \tag{2-1}$$

(A2) (uniform ellipticity) For all $y \in \mathbb{T}^d$, the matrix $A(y) = (a_{ij}(y))$ is uniformly elliptic in the sense that, for all $\xi \in \mathbb{R}^d$, one has

$$\lambda|\xi|^2 \le \xi^T A(y)\xi = \sum_{i,j=1}^{d} \xi_i a_{ij}(y)\xi_j \le \Lambda|\xi|^2. \tag{2-2}$$

(A3) (smoothness) For some γ, M with $\gamma \in (0,1]$ and $M > 0$, one has

$$\|A\|_{C^\gamma(\mathbb{T}^d)} \le M. \tag{2-3}$$

We henceforth refer to the above assumptions together as (A).

Random potential. For the random field $q(x,\omega)$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we assume:

(P) (stationarity and ergodicity) There exists an ergodic group of $\mathbb{P}$-preserving transformations $(\tau_x)_{x\in\mathbb{R}^d}$ on Ω, where ergodicity means that $E \in \mathcal{F}$ and

$$\tau_x E = E \quad \text{for all } x \in \mathbb{R}^d$$

imply that $\mathbb{P}(E) \in \{0,1\}$. The random potential $q(y,\omega)$ is given by $\tilde{q}(\tau_y\omega)$, where $\tilde{q} : \Omega \to \mathbb{R}$ is a random variable satisfying

$$0 \le \tilde{q}(\omega) \le M \quad \text{for all } \omega \in \Omega. \tag{2-4}$$

Further assumptions on q. The above assumptions are sufficient for proving the homogenization result. However, to estimate the size of the homogenization error and to characterize the limiting distribution of the random fluctuation, more assumptions on the random field $q(\cdot,\omega)$ are necessary.

To simplify notations, we write in the sequel

$$q(x,\omega)=\bar{q}+\nu(x,\omega),$$

where $\bar{q}$ is the mean of q and ν is the fluctuation. Note that $\bar{q}$ is a deterministic constant and ν is a mean zero stationary ergodic random field. The autocorrelation function $R(x)$ of q (and hence ν) is defined as

$$R(x)=\mathbb{E}\big(\nu(x+y,\omega)\nu(y,\omega)\big),\quad \sigma^2:=\int_{\mathbb{R}^d}R(x)\,dx. \tag{2-5}$$

By Bochner's theorem, $R(x)$ is a positive definite function and $\sigma^2\ge 0$. We assume that $\sigma>0$. When R is integrable on $\mathbb{R}^d$, i.e., $\sigma^2<\infty$, we say that q has short-range correlations; we say q has long-range correlations if otherwise. We state and prove the main results in the setting where q has short-range correlations, and mention the corresponding results for the long-range correlation setting in Section 6.

Short-range correlated random fields. In this case, we make an assumption on the rate of decay of the correlation function. We denote by $\mathcal{C}$ the set of compact sets in $\mathbb{R}^d$, and for two sets K_1, K_2 in $\mathcal{C}$, the distance $d(K_1,K_2)$ is defined to be

$$d(K_1,K_2)=\min_{x\in K_1,y\in K_2}|x-y|.$$

Given any compact set $K\subset\mathcal{C}$, we denote by $\mathcal{F}_K$ the σ-algebra generated by the random variables $\{q(x):x\in K\}$. We define the "maximal correlation coefficient" ϱ of q as follows: for each $r>0$, $\varrho(r)$ is the smallest value such that the bound

$$\mathbb{E}\big(\varphi_1(q)\varphi_2(q)\big)\le\varrho(r)\sqrt{\mathbb{E}(\varphi_1^2(q))\,\mathbb{E}(\varphi_2^2(q))} \tag{2-6}$$

holds for any two compact sets $K_1,K_2\in\mathcal{C}$ such that $d(K_1,K_2)\ge r$ and for any two random variables of the form $\varphi_i(q)$, with $i=1,2$, such that $\varphi_i(q)$ is $\mathcal{F}_{K_i}$-measurable and $\mathbb{E}\varphi_i(q)=0$. We assume that

(S) The maximal correlation function satisfies $\varrho^{1/2}\in L^1(\mathbb{R}_+,r^{d-1}dr)$; that is,

$$\int_0^\infty \varrho^{\frac{1}{2}}(r)r^{d-1}\,dr<\infty.$$

Assumptions on the mixing coefficient ϱ of random media have been used in [Bal 2008; Bal and Jing 2011; Hairer et al. 2013]; we refer to these papers for explicit examples of random fields satisfying the assumptions. We note that the autocorrelation function $R(x)$ can be bounded by ϱ. For any $x\in\mathbb{R}^d$,

$$|R(x)|=\big|\mathbb{E}(q(x)-\mathbb{E}q)(q(0)-\mathbb{E}q)\big|\le\varrho(|x|)\,\mathrm{Var}(q).$$

By (2-4), q, and hence its variance, is bounded. In view of (S) and the fact that one can assume $\varrho\in[0,1]$ (hence $\varrho\le\sqrt{\varrho}$), we find that R is integrable. Therefore, (S) implies that $q(x,\omega)$ has short-range correlations. In fact, (S) is a much stronger assumption, and not necessary for the main results of this paper to hold. In Section 7, we will provide alternative and less restrictive assumptions that are sufficient.

However, using the assumption (S) and Lemma 4.3 below, we can simplify significantly certain fourth-order moment estimates of the random potential $\nu(x,\omega)$; such estimates appear often in the study of the limiting distribution of the homogenization error.

Notations. Throughout the paper, by *universal parameters* we refer to λ, Λ, γ and M in the assumptions (A), the autocorrelation function R, σ^2, and the mixing coefficients ϱ, the domain D and its boundary ∂D, and the dimension d. If a constant C depends only on these parameters, we say either C depends on universal parameters or C is a universal constant. For the random potential $\nu(x,\omega)$ and the functions $\varrho(x)$, $R(x)$, etc. which are related to ν, we use $\nu^\varepsilon, \varrho^\varepsilon, R^\varepsilon$, etc. to denote the scaled versions. For instance, $\nu^\varepsilon(x,\omega)$ is shorthand notation for $\nu\big(\frac{x}{\varepsilon}\big)$. We use the notation $H^s(K)$, with $s \ge 0$, for the Sobolev or the fractional Sobolev space $W^{s,2}(K)$ on some domain $K \subset \mathbb{R}^d$; when K is bounded, we use $H_0^s(K)$ for the subspace that consists of functions having trace zero at ∂K; note that $H_0^s(\mathbb{R}^d) = H^s(\mathbb{R}^d)$. We denote by $H^{-s}(K)$, with $s > 0$, the dual space $(H_0^s(K))'$. For any Hilbert space $\mathcal{H}$, we denote the inner product in $\mathcal{H}$ by $(\cdot\,,\cdot)_{\mathcal{H}}$; when $\mathcal{H} = L^2(D)$, we very often omit the subscript and write $(\cdot\,,\cdot)$ instead. We use $\langle f, g\rangle$ whenever the formal integral $\int_D fg$ makes sense. We typically use $\mathbb{1}_A$ for the indication function of a set $A \subset \mathbb{R}^d$, or if A is a statement, the indication function of A being true. Finally, for two real numbers a and b, we use $a \wedge b$ as a shorthand notation for $\min\{a,b\}$, and $a \vee b$ means $\max\{a,b\}$.

2B. ***Probability distribution on functional spaces.*** We view the random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$ in the homogenization error as random elements in certain functional spaces, and aim to find the limit of its law in that space. It turns out that the choice of functional spaces depends on the spatial dimension d.

When $d = 1$, one can choose the space $C(D)$ of continuous functions. In fact, convergence in distribution in $C(D)$ was proved in [Bal 2008] for random diffusion coefficient $a(x,\omega)$ with random potential $q(x,\omega)$, both having short-range correlations. In this paper, we prove that for $d = 2, 3$, the space can be chosen as $L^2(D)$ and for $d = 4, 5$, the space can be chosen as $H^{-1}(D)$. Note that both choices are Hilbert spaces. We recall some facts concerning weak convergence of probability measures on Hilbert spaces. We refer to the books of Billingsley [1999] and Parthasarathy [1967] for more details.

Probability distributions on a Hilbert space. Let $\mathcal{H}$ be a separable Hilbert space, and let $X(\omega)$ be an $\mathcal{H}$-valued random element on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then X determines a probability measure P^X on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$, where $\mathcal{B}(\mathcal{H})$ denotes the Borel σ-algebra generated by open sets in $\mathcal{H}$, by

$$P^X(\mathcal{S}) = \mathbb{P}(X \in \mathcal{S}) \quad \text{for any } \mathcal{S} \in \mathcal{B}(\mathcal{H}). \tag{2-7}$$

We say a family $\{X^\varepsilon\}_{\varepsilon\in(0,1)}$ of random elements in $\mathcal{H}$ converges in probability distribution (or in law), as $\varepsilon \to 0$, to another random element X on $\mathcal{H}$, if the probability measures P^{X^ε} converge weakly to P^X; i.e., for any real bounded continuous functional $f : \mathcal{H} \to \mathbb{R}$,

$$\int_{\mathcal{H}} f(g)\, dP^{X^\varepsilon}(g) \to \int_{\mathcal{H}} f(g)\, dP^X(g).$$

In particular, any probability measure P on a separable Hilbert space $\mathcal{H}$ is determined by its characteristic function $\phi^P : \mathcal{H} \to \mathbb{C}$,

$$\phi^P(h) = \int_{\mathcal{H}} e^{i(h,g)_{\mathcal{H}}}\, dP(g). \tag{2-8}$$

Moreover, the following result holds:

Theorem 2.1 [Parthasarathy 1967, Chapter VI, Lemma 2.1]. *Let $\{X^\varepsilon\}_{\varepsilon\in(0,1)}$ and X be random elements in $\mathcal{H}$, possibly defined on different probability spaces. Then X^ε converges to X in law in $\mathcal{H}$, as $\varepsilon \to 0$, if the family of probability measures $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ is tight and for any $h \in \mathcal{H}$,*

$$\lim_{\varepsilon\to 0} \phi^{P^{X^\varepsilon}}(h) = \phi^{P^X}(h). \tag{2-9}$$

Remark 2.2. Let $\mathcal{H} = L^2(D)$, which is a separable Hilbert space, and let X be a random element in $L^2(D)$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The characteristic function of P^X can be calculated as follows: for any $h \in L^2(D)$,

$$\phi^{P^X}(h) = \int_{\mathbb{R}} e^{iz}\, dP^X\big(\{(h, g) > z\}\big) = \int_{\mathbb{R}} e^{iz}\, d\mathbb{P}\big(\{(h, X(\omega)) > z\}\big) = \mathbb{E}\, e^{i(h,X)}. \tag{2-10}$$

Therefore, to prove that X^ε converges in distribution to X as L^2-paths, it suffices to show that $\{P^{X^\varepsilon}\}$ is tight and that for any $h \in L^2(D)$,

$$(h, X^\varepsilon) \xrightarrow{\text{distribution}} (h, X); \tag{2-11}$$

that is, the random variables (h, X^ε) converge in distribution to the random variable (h, X).

In Theorem A.1 in the Appendix, we provide a tightness criterion for $\{P^{X^\varepsilon}\}$ on $L^2(D)$, with the assumption that $\{X^\varepsilon(\cdot, \omega)\}$ is in $H_0^s(D)$ for certain $s > 0$. The criterion is sufficient but by no means necessary. Nevertheless, it is very handy for our analysis since the random fields X^ε that we are dealing with come from solutions of (1-1), and hence are naturally in $H_0^s(D)$.

2C. ***Main results.*** We now state the main results of the paper under the assumption that $q(x, \omega)$ has *short-range correlations*. Analogous results for the long-range correlation setting will be presented in Section 6.

The first main theorem concerns how the homogenization error scales.

Theorem 2.3. *Let $D \subset \mathbb{R}^d$ be an open bounded $C^{1,1}$ domain, u^ε and u be the solutions to* (1-1) *and* (1-2) *respectively. Suppose that* (A), (P) *and* (S) *hold, $f \in L^2(D)$ and $2 \le d \le 7$. Then, there exists positive constant C, depending only on the universal parameters, such that*

$$E\, \|u^\varepsilon - u\|_{L^2} \le C\varepsilon \|f\|_{L^2}. \tag{2-12}$$

Moreover,

$$E\, \|u^\varepsilon - \mathbb{E}u^\varepsilon\|_{L^2} \le \begin{cases} C\varepsilon^{2\wedge\frac{d}{2}} \|f\|_{L^2} & \text{if } d \ne 4, \\ C\varepsilon^2 |\log\varepsilon|^{\frac{1}{2}} \|f\|_{L^2} & \text{if } d = 4. \end{cases} \tag{2-13}$$

Furthermore, for any $\varphi \in L^2(D)$,

$$\mathbb{E}\, \big|(u^\varepsilon - \mathbb{E}u^\varepsilon, \varphi)_{L^2}\big| \le C\varepsilon^{\frac{d}{2}} \|\varphi\|_{L^2} \|f\|_{L^2}. \tag{2-14}$$

This theorem provides $L^1(\Omega, L^2(D))$-estimates of $u^\varepsilon - u$ and its random part, and its proof is detailed in Section 4. We note that the size of the full homogenization error is much larger than that of its random part. This is because the oscillations in the diffusion coefficients cause some deterministic fluctuation in the solution of size $O(\varepsilon)$, as in standard periodic homogenization. The additional random fluctuation caused by the short-range correlated random potential scales like $\varepsilon^{(d\wedge 4)/2}$ in the energy norm, and scales like $\varepsilon^{d/2}$ in the weak topology. These results agree with the case of nonoscillatory diffusion coefficients; see [Bal 2008; Bal and Jing 2011]. The next result exhibits the limiting law of the rescaled random fluctuation $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}u^\varepsilon)$.

Theorem 2.4. *Suppose the assumptions in Theorem 2.3 hold. Let σ be defined as in* (2-5) *and $G(x, y)$ be the Green's function of* (1-2). *Let $W(y)$ denote the standard d-parameter Wiener process. Then*

(i) *For $d = 2, 3$, as $\varepsilon \to 0$,*
$$\frac{u^\varepsilon - \mathbb{E}u^\varepsilon}{\sqrt{\varepsilon^d}} \xrightarrow{\text{distribution}} \sigma \int_D G(x, y)u(y)\, dW(y) \quad \text{in } L^2(D). \tag{2-15}$$

(ii) *For $d = 4, 5$, as $\varepsilon \to 0$, the above holds as convergence in law in $H^{-1}(D)$.*

The proof of item (i) above can be found on page 217 and that of item (ii) is on page 220.

Remark 2.5. The integral on the right-hand side of (2-15) is understood, for each fixed x, as a Wiener integral in y with respect to the multiparameter Wiener process $W(y)$. Let X denote the result. For $d = 2, 3$, because the Green's function $G(x, y)$ is square integrable, X is a random element in $L^2(D)$. For $d = 4, 5$, X is understood through the Fourier transform of its distribution: given $h^* \in H^{-1}(D)$, $\phi^{P^X}(h^*)$ is defined to be $\boldsymbol{E}\exp\big(i\sigma \int_D \langle G(\cdot\,, y), h^*(\cdot)\rangle u(y)\, dW(y)\big)$, where $\boldsymbol{E}$ is the expectation with respect to the law of W.

Remark 2.6. We expect that the scaling factor for the random fluctuation, with respect to the weak topology, should be $\varepsilon^{-d/2}$ in all dimensions. More precisely, for any $\varphi \in L^2(D)$, we expect that $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}u^\varepsilon, \varphi)$ should converge in distribution for all dimensions. However, in this paper we control this term only for $d \le 7$. This constraint is not intrinsic, and is mainly due to the fact that we stopped at second order iteration in the series expansion (4-11). In fact, if higher- (than six or more) order moments of the random field are under control, we can iterate as many times as we need in (4-11) until the last term is small, and use higher-order moments to estimate the terms in between; see Remark 4.6 below.

The spatial dimension plays an intrinsic role on the choice of topology that one should use for the limiting distribution of the random fluctuations. Indeed, for the term $X^\varepsilon = -\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon$ to converge in law in $L^2(D)$, it is necessary that $\mathbb{E}\|X^\varepsilon\|_{L^2}^2$ is controlled uniformly in ε. In view of the singularity of the Green's function, namely, of order $|x-y|^{-d+2}$ near the diagonal, we expect to control $\mathbb{E}\|X^\varepsilon\|_{L^2}^2$ only for $d \le 3$, and similarly, we expect to have convergence in law in H^{-1} only for $d < 6$. Nevertheless, we expect that convergence in law in $H^{-k}(D)$, for certain $k > 0$ increasing with respect to d, could be proved, provided more controls on the random field are available.

Finally, we remark that other topologies, e.g., those in [Bal et al. 2012; Gu and Mourrat 2015], can be considered for the law of the random fluctuation as well. In particular, tightness criteria in the Hölder space C^α, with α possibly negative, were established in [Mourrat 2015]. By a formal scaling argument, the short-range noise ν_ε belongs to the Hölder class C^{0-} and the Green's function is in C^{2-d}. The convergence of X^ε, which is essentially a convolution of the Green's function with the noise and then divided by $\varepsilon^{d/2}$, should take place in C^α, for $\alpha < -\frac{d}{2} + 2$. In fact, this agrees with the constraint that convergence in L^2 can be expected only for $d < 4$, and convergence in H^{-1} for $d < 6$. It would be interesting to pursue this direction of studies further.

3. Homogenization and periodic error estimates

The following homogenization result for (1-1), without the random potential $q^\varepsilon(x,\omega)$, is well known. The effect of the presence of q^ε turns out to be minor for homogenization; nevertheless, we include a proof here for the sake of completeness.

Theorem 3.1. *Assume* (A1), (A2) *and* (P) *hold. Then there exists $\Omega_1 \in \mathcal{F}$ such that $\mathbb{P}(\Omega_1) = 1$, and for all $\omega \in \Omega_1$, the solution u^ε of* (1-1) *converges to the solution u of* (1-2) *weakly in $H^1(D)$ and strongly in $L^2(D)$ for any $f \in H^{-1}(D)$.*

Let $\mathcal{L}_\varepsilon$ denote the differential operator

$$-\frac{\partial}{\partial x_i}\Big(a_{ij}\Big(\frac{x}{\varepsilon}\Big)\frac{\partial}{\partial x_j}\Big) + \bar{q}, \tag{3-1}$$

and let $\mathcal{L}^{\varepsilon,\omega}$ be the differential operator $\mathcal{L}_\varepsilon + \nu(\frac{x}{\varepsilon},\omega)$. We remark that $\mathcal{L}_\varepsilon$ has highly oscillatory but deterministic coefficients while $\mathcal{L}^{\varepsilon,\omega}$ has, in addition, a highly oscillatory and random potential. Let $\mathcal{G}^{\varepsilon,\omega}$ and $\mathcal{G}_\varepsilon$ be the solution operator of the Dirichlet boundary problems associated to $\mathcal{L}^{\varepsilon,\omega}$ and $\mathcal{L}_\varepsilon$. Owing to the conditions (2-2) and (2-4), $\mathcal{G}^{\varepsilon,\omega}$ is well-defined for any $\omega \in \Omega$. Moreover, we have the standard estimate, for any $\omega \in \Omega$ and $\varepsilon > 0$,

$$\|\mathcal{G}^{\varepsilon,\omega} f\|_{H^1(D)} \le C\|f\|_{H^{-1}(D)}, \tag{3-2}$$

with some constant C that depends on the universal parameters, and neither on ω nor ε. By the same token, $\mathcal{G}_\varepsilon$ is well-defined and shares the same estimate above.

Proof of Theorem 3.1. Step 1: For each $\omega \in \Omega$, the solution u^ε of (1-1) is given by $\mathcal{G}^{\varepsilon,\omega} f$, which satisfies the standard estimates

$$\|u^\varepsilon\|_{H^1(D)} + \|A^\varepsilon \nabla u^\varepsilon\|_{L^2(D)} + \|q^\varepsilon(x,\omega)u^\varepsilon\|_{L^2(D)} \le C,$$

where C depends the universal parameters and f and is uniform in ε and ω. As a result, due to the compact embeddings $H^1(D) \hookrightarrow L^2(D) \hookrightarrow H^{-1}(D)$, through a subsequence $\varepsilon_j(\omega) \to 0$, which by an abuse of notation is still denoted by ε, we have

$$\begin{aligned} \nabla u^\varepsilon(\cdot,\omega) &\xrightarrow[\varepsilon\to 0]{L^2} \nabla v(\cdot,\omega), & A\Big(\frac{\cdot}{\varepsilon}\Big)\nabla u^\varepsilon(\cdot,\omega) &\xrightarrow[\varepsilon\to 0]{L^2} \xi(\cdot,\omega), \\ u^\varepsilon(\cdot,\omega) &\xrightarrow[\varepsilon\to 0]{L^2} v(\cdot,\omega), & q\Big(\frac{\cdot}{\varepsilon},\omega\Big)u^\varepsilon(\cdot,\omega) &\xrightarrow[\varepsilon\to 0]{H^{-1}} p(\cdot,\omega) \end{aligned} \tag{3-3}$$

for some function $v(\cdot,\omega) \in H^1(D)$ and some vector-valued function $\xi(\cdot,\omega) \in [L^2(D)]^d$.

Step 2: Recall that $\{\chi^k\}_{k=1}^d$ are the correctors defined in (1-4), and we can extend them periodically to functions defined on $\mathbb{R}^d$. Since $A(y)(e_k + \nabla\chi^k(y))$ is periodic, we have that

$$A\Big(\frac{x}{\varepsilon}\Big)\Big(e_k + (\nabla\chi^k)\Big(\frac{x}{\varepsilon}\Big)\Big) \xrightarrow{L^2} \int_{\mathbb{T}^d} A(y)\big(e_k(y) + \nabla\chi^k(y)\big)\,dy = \bar{A}e_k. \tag{3-4}$$

For the same reason and the fact $\int_{\mathbb{T}^d} \nabla\chi^k \, dy = 0$, we have

$$e_k + (\nabla\chi^k)\Big(\frac{x}{\varepsilon}\Big) \xrightarrow{L^2} \int_{\mathbb{T}^d} e_k + \nabla\chi^k(y)\, dy = e_k. \tag{3-5}$$

Now fix an arbitrary function $\varphi \in C_0^\infty(D)$. For each fixed $\omega \in \Omega$, let $\varepsilon(\omega) \to 0$ be the subsequence in Step 1. Consider the integral

$$\int_D A\Big(\frac{x}{\varepsilon}\Big)\nabla u^\varepsilon(x,\omega)\cdot\nabla\Big\{x_k + \varepsilon\chi^k\Big(\frac{x}{\varepsilon}\Big)\Big\}\varphi(x)\, dx.$$

On one hand, in view of the third item in (3-3), (3-5), and the facts that $\operatorname{div}(A^\varepsilon\nabla u^\varepsilon) = -f + q^\varepsilon u^\varepsilon$ converges in H^{-1} (to $-f + p(\cdot,\omega)$, where p is defined in (3-3)) and that $e_k + (\nabla\chi^k)(x/\varepsilon)$ is curl-free, by the div–curl lemma [Jikov et al. 1994, Lemma 1.1], the above integral satisfies

$$\lim_{\varepsilon\to 0}\int_D A\Big(\frac{x}{\varepsilon}\Big)\nabla u^\varepsilon(x,\omega)\cdot\nabla\Big\{x_k + \varepsilon\chi^k\Big(\frac{x}{\varepsilon}\Big)\Big\}\varphi(x)\, dx = \int_D \xi(x,\omega)\cdot e_k\varphi(x)\, dx.$$

On the other hand, in view of the first item in (3-3), (3-4), and the facts that $\operatorname{div}(A^\varepsilon(e_k + \nabla\chi^k(x/\varepsilon)))$ converges in H^{-1} (they are all equal to zero) and that ∇u^ε is curl-free, by the div–curl lemma, we have

$$\lim_{\varepsilon\to 0}\int_D A\Big(\frac{x}{\varepsilon}\Big)\nabla u^\varepsilon(x,\omega)\cdot\nabla\Big\{x_k + \varepsilon\chi^k\Big(\frac{x}{\varepsilon}\Big)\Big\}\varphi(x)\, dx = \int_D \nabla v\bar{A}\cdot e_k\varphi(x)\, dx.$$

The two limits above must be equal, and it follows that $\xi(\cdot,\omega) = \bar{A}\nabla v(\cdot,\omega)$ in distribution.

Step 3: Recall that the stationary random potential $q(x,\omega)$ can be written as $\tilde{q}(\tau_x\omega)$, where $\tilde{q}$ is an essentially bounded random variable on Ω. By the Birkhoff ergodic theorem [Jikov et al. 1994, Theorem 7.2], there exists $\Omega_1 \in \mathcal{F}$ with $\mathbb{P}(\Omega_1) = 1$, and for each $\omega \in \Omega_1$,

$$q\Big(\frac{x}{\varepsilon},\omega\Big) = \tilde{q}(\tau_{\frac{x}{\varepsilon}}\omega) \xrightarrow{L^\alpha_{\mathrm{loc}}(\mathbb{R}^d)} \bar{q} = \mathbb{E}q(0,\omega) \tag{3-6}$$

for any $\alpha \in (1,\infty)$. From the weak formulation of u^ε, for any $\omega \in \Omega_1$ and for any $\varphi \in C_0^\infty(D)$, we have

$$\int_D A\Big(\frac{x}{\varepsilon}\Big)\nabla u^\varepsilon(x,\omega)\cdot\nabla\varphi(x)\, dx + \int_D q\Big(\frac{x}{\varepsilon},\omega\Big)u^\varepsilon(x)\varphi(x)\, dx = \int_D f(x)\varphi(x)\, dx.$$

Passing to the limit along the subsequence $\varepsilon(\omega)$ found in Step 1, we have

$$\int_D \bar{A}\nabla v\cdot\nabla\varphi + \int_D \bar{q}v(x)\varphi(x)\, dx = \int_D f(x)\varphi(x)\, dx - \lim_{\varepsilon\to 0}\int_D q\Big(\frac{x}{\varepsilon},\omega\Big)(u^\varepsilon - v)\varphi(x)\, dx.$$

The first term on the left follows from (3-3) and the fact that $\xi = \bar{A}\nabla v$; the second term on the left is due to (3-6). Finally, the last term on the right-hand side is zero since q is uniformly bounded and $u^\varepsilon - v$ converges to zero strongly in $L^2(D)$. Consequently, the above limit shows that v solves the homogenized equation (1-2). By uniqueness of the homogenized problem, we must have that $v = u$ and v is deterministic.

Finally, for each $\omega \in \Omega_1$, by the weak compactness in $H^1(D)$ and the uniqueness of the possible limit, the whole sequence u^ε converges to u. This proves the homogenization theorem. □

Remark 3.2. We remark that the same proof works in the case when (a_{ij}) is not symmetric; indeed, it suffices to replace χ^k above by the solution of the adjoint corrector equation. The same idea of proof can also be carried out in the case when (a_{ij}) are stationary ergodic random fields; indeed, the corrector equation in that case is much more involved but, by now, its solution and analogs of (3-4) and (3-5), are well known.

3A. ***Decomposition of the homogenization error.*** To separate the fluctuations in the homogenization error $u^\varepsilon - u$ that are due to the periodic oscillations in the diffusion coefficients from those due to the random potential, we introduce the function v^ε which solves the following deterministic problem:

$$\begin{cases} -\dfrac{\partial}{\partial x_i}\Big(a_{ij}\Big(\dfrac{x}{\varepsilon}\Big)\dfrac{\partial}{\partial x_j}v^\varepsilon(x,\omega)\Big) + \bar{q}v^\varepsilon(x,\omega) = f(x), & x \in D, \\ v^\varepsilon(x) = 0, & x \in \partial D. \end{cases} \tag{3-7}$$

Here, the potential field is already homogenized, and we expect that $v^\varepsilon - u$ filters out the effect of the random potential. The problem above is well-posed and its solution v^ε is given by $\mathcal{G}_\varepsilon f$.

The standard periodic homogenization theory yields that v^ε converges weakly in $H^1(D)$ and strongly in $L^2(D)$ to u for any $f \in H^{-1}(D)$. Using this function, we can write the homogenization error for (1-1) as

$$u^\varepsilon - u = (u^\varepsilon - v^\varepsilon) + (v^\varepsilon - u). \tag{3-8}$$

The deterministic part of the homogenization error is

$$\mathbb{E}u^\varepsilon - u = \mathbb{E}(u^\varepsilon - v^\varepsilon) + (v^\varepsilon - u), \tag{3-9}$$

and the random fluctuation part of the homogenization error is

$$u^\varepsilon - \mathbb{E}u^\varepsilon = (u^\varepsilon - v^\varepsilon) - \mathbb{E}(u^\varepsilon - v^\varepsilon). \tag{3-10}$$

The deterministic part of the homogenization error hence contains two parts, the mean of $u^\varepsilon - v^\varepsilon$ and the periodic homogenization error $v^\varepsilon - u$. Estimates for the second part amount to the convergence rate of periodic homogenization, and we recall some of the well-known results below, together with uniform-in-ε estimates of the Green's function associated to $\mathcal{G}_\varepsilon$. We postpone the estimates for $\mathbb{E}(u^\varepsilon - v^\varepsilon)$ to the next section.

Theorem 3.3 (estimates in periodic homogenization). *Let $D \subset \mathbb{R}^d$ be an open bounded $C^{1,1}$-domain, and v^ε and u be the solutions to* (3-7) *and* (1-2) *respectively. Let $G_\varepsilon(x, y)$, with $x, y \in D$, be the Green's function associated to the Dirichlet problem of* (3-7). *Assume* (A) *holds. Then there exists positive constant C, depending only on the universal parameters, such that*

(i) *for any $f \in L^2(D)$, we have $\|v^\varepsilon - u\|_{L^2} \le C\varepsilon\|f\|_{L^2}$,*

(ii) *for $d \ge 2$ and for any $x, y \in D$, $x \ne y$, we have that $G_\varepsilon(x, y)$ satisfies*

$$|G_\varepsilon(x,y)| \le \begin{cases} C|x-y|^{2-d} & \text{if } d \ne 2, \\ C\big(1 + \big|\log|x-y|\big|\big) & \text{if } d = 2, \end{cases} \tag{3-11}$$

and

$$|\nabla G_\varepsilon(x,y)| \le C|x-y|^{1-d}. \tag{3-12}$$

The $O(\varepsilon)$-error estimates in L^2 were proved in [Moskow and Vogelius 1997] for $d = 2$, and in [Griso 2006] for general $C^{1,1}$-domains; see also [Kenig et al. 2012]. The uniform-in-ε estimates on the Green's function and its gradient can be found, e.g., in [Avellaneda and Lin 1987; 1991; 2015]. In particular, (3-11) was proved in [Avellaneda and Lin 1987, Theorem 13]; the estimate (3-12) follows from an interior Lipschitz estimate, e.g., [Avellaneda and Lin 1987, Lemma 16], if the distance between x and y is smaller compared with their distance from the boundary, and it follows from a boundary Lipschitz estimate, e.g., [Avellaneda and Lin 1987, Lemma 20], if otherwise; see also the proof of [Armstrong and Shen 2015, Theorem 1.1].

The homogeneities in these bounds are the same as those for the Green's function associated to constant coefficient equations, namely the Laplace equations. The striking fact that these bounds still hold for oscillatory equations is due to the fact that the problem (3-7) homogenizes to constant (smooth) coefficient equations. Periodicity or other structural assumptions on the coefficients are crucial. We remark also that it is to obtain such pointwise estimates that the Hölder regularity of the diffusion matrix, i.e., assumption (A3), is needed.

4. Estimates for the homogenization error

In this section, we estimate the size of the homogenization error $u^\varepsilon - u$. In view of the decomposition (3-8), (3-9), (3-10) and the error estimates in Theorem 3.3, it suffices to focus on the intermediate homogenization error $u^\varepsilon - v^\varepsilon$, with $v^\varepsilon = \mathcal{G}_\varepsilon f$ defined in (3-7).

We introduce the function w^ε which solves

$$\mathcal{L}_\varepsilon w^\varepsilon = -\nu_\varepsilon v^\varepsilon, \tag{4-1}$$

with homogeneous Dirichlet boundary condition. With the notations $\mathcal{G}_\varepsilon$ and $\mathcal{G}^{\varepsilon,\omega}$ introduced earlier, w^ε is given by $-\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon$. It follows that

$$\mathcal{L}^{\varepsilon,\omega}(u^\varepsilon - v^\varepsilon - w^\varepsilon) = -\nu_\varepsilon w^\varepsilon,$$

and $u^\varepsilon - v^\varepsilon - w^\varepsilon$ vanishes at the boundary. Hence we have $u^\varepsilon - v^\varepsilon - w^\varepsilon = -\mathcal{G}^{\varepsilon,\omega}\nu_\varepsilon w^\varepsilon$. Due to the assumptions (A), $\mathcal{G}^{\varepsilon,\omega}$ is uniformly (in ε and ω) bounded as a linear operator $L^2 \to L^2$; we have

$$\|u^\varepsilon - v^\varepsilon\|_{L^2} \le C\|w^\varepsilon\|_{L^2}. \tag{4-2}$$

An estimate of $u^\varepsilon - v^\varepsilon$ thus follows from this result:

Lemma 4.1. *Let $v^\varepsilon = \mathcal{G}_\varepsilon f$ and w^ε be as above. Under the same conditions of Theorem 2.3, there exists a universal constant C and*

$$\mathbb{E}\|w^\varepsilon\|^2_{L^2(D)} \le \begin{cases} C\varepsilon^{d\wedge 4}\|f\|^2_{L^2} & \text{if } d \ne 4, \\ C\varepsilon^4|\log\varepsilon|\,\|f\|^2_{L^2} & \text{if } d = 4. \end{cases} \tag{4-3}$$

Proof. Using the Green's function G_ε, we write

$$w^\varepsilon(x,\omega) = \int_D G_\varepsilon(x,y)\nu\left(\frac{y}{\varepsilon}\right)v^\varepsilon(y)\,dy. \tag{4-4}$$

The L^2-energy of w^ε is then

$$\|w^\varepsilon(\cdot,\omega)\|^2_{L^2} = \int_{D^3} G_\varepsilon(x,y)G_\varepsilon(x,z)\nu\left(\frac{y}{\varepsilon}\right)\nu\left(\frac{z}{\varepsilon}\right)v^\varepsilon(y)v^\varepsilon(z)\,dy\,dz\,dx.$$

Taking the expectation and using the definition of the autocorrelation function R of q, we have

$$\mathbb{E}\|w^\varepsilon(\cdot,\omega)\|_{L^2}^2 = \int_{D^3} G_\varepsilon(x,y)G_\varepsilon(x,z)R\Big(\frac{y-z}{\varepsilon}\Big)v^\varepsilon(y)v^\varepsilon(z)\,dy\,dz\,dx. \tag{4-5}$$

Integrate over x first. Apply the uniform estimates (3-11) and the fact (see, e.g., Lemma A.1 of [Bal and Jing 2011]): for any $y \neq z$, $0 < \alpha, \beta < d$,

$$\int_D \frac{dx}{|x-y|^{d-\alpha}|x-z|^{d-\beta}} \le \begin{cases} C & \text{if } \alpha+\beta > d, \\ C\big(1+\big|\log|y-z|\big|\big) & \text{if } \alpha+\beta = d, \\ C|y-z|^{\alpha+\beta-d} & \text{if } \alpha+\beta < d. \end{cases} \tag{4-6}$$

We get

$$\int_D |G_\varepsilon(x,y)G_\varepsilon(x,z)|\,dx \le \begin{cases} C|y-z|^{-((d-4)\wedge 0)} & \text{if } d \neq 4, \\ C(1+\log|y-z|) & \text{if } d = 4. \end{cases} \tag{4-7}$$

Hence, if $d \ge 2$ and $d \neq 4$,

$$\mathbb{E}\|w^\varepsilon(\cdot,\omega)\|_{L^2}^2 \le C\int_{D^2} \frac{|v^\varepsilon(y)v^\varepsilon(z)|}{|y-z|^{(d-4)\vee 0}}\Big|R\Big(\frac{y-z}{\varepsilon}\Big)\Big|\,dy\,dz\,dx.$$

When $d = 4$, the term $(|y-z|^{(d-4)\vee 0})^{-1}$ should be replaced by $1+\big|\log|y-z|\big|$. In any case, the above yields a bound of the form

$$\mathbb{E}\|w^\varepsilon(\cdot,\omega)\|_{L^2}^2 \le C\int_{\mathbb{R}^d} |\tilde{v}^\varepsilon(y)|(K^\varepsilon * \tilde{v}^\varepsilon)(y)\,dy. \tag{4-8}$$

Here, $\tilde{v}^\varepsilon = v^\varepsilon \mathbb{1}_D$ and $\mathbb{1}_D$ denotes the indicator function of the set D, $K^\varepsilon(y) = R\big(\frac{y}{\varepsilon}\big)|y|^{(4-d)\wedge 0}\mathbb{1}_{B_\rho}(y)$ if $d \neq 4$ and $K^\varepsilon(y) = R\big(\frac{y}{\varepsilon}\big)\big(1+\mathbb{1}_{B_\rho}(y)\big|\log|y|\big|\big)$ if $d = 4$. Here, B_ρ is the ball centered at zero with radius ρ and ρ is the diameter of D. We check that, when $d \neq 4$,

$$\|K^\varepsilon(y)\|_{L^1} \le \int_{\mathbb{R}^d}\Big|R\Big(\frac{y}{\varepsilon}\Big)\Big|\frac{1}{|y|^{(d-4)\vee 0}}\,dy = \frac{\varepsilon^d}{\varepsilon^{(d-4)\vee 0}}\int_{\mathbb{R}^d}\frac{|R(y)|}{|y|^{(d-4)\vee 0}} = C\varepsilon^{d\wedge 4}, \tag{4-9}$$

where in the last inequality we used $R \in L^\infty \cap L^1(\mathbb{R}^d)$. Similarly, when $d = 4$,

$$\|K^\varepsilon(y)\|_{L^1} = \int_{B_\rho}\Big|R\Big(\frac{y}{\varepsilon}\Big)\Big|\big(1+\big|\log|y|\big|\big)\,dy = \varepsilon^4\int_{B_{\rho/\varepsilon}}|R(y)|\big(1+\big|\log|\varepsilon y|\big|\big) \le C\varepsilon^4|\log\varepsilon|. \tag{4-10}$$

To get the last inequality, we evaluate the integral on B_1 and $B_{\rho/\varepsilon}\setminus B_1$, and bound $\big|\log|\varepsilon y|\big|$ by $\big|\log|\varepsilon\rho|\big|$ for the second part. Applying Hölder's and then Young's inequalities to (4-8), we get

$$\mathbb{E}\|w^\varepsilon\|_{L^2}^2 \le C\|K^\varepsilon\|_{L^1}\|v^\varepsilon\|_{L^2}^2 \le C\|K^\varepsilon\|_{L^1}\|f\|_{L^2}^2.$$

Combining this with the estimates in (4-9) and (4-10), we complete the proof of the lemma. ☐

4A. ***Scaling of the energy in the random fluctuation.*** Now we estimate the $L^2(D)$-norm (the energy) of the random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$ which, in view of (3-10), is the same as the fluctuation $u^\varepsilon - v^\varepsilon - \mathbb{E}(u^\varepsilon - v^\varepsilon)$.

Using the first-order corrector w^ε defined by (4-1), and following the approach of [Bal 2008; Bal and Jing 2011], we can derive an expansion formula for $u^\varepsilon - v^\varepsilon$ as follows. Rewrite the equations (1-1) and (3-7) as

$$\mathcal{L}_\varepsilon u^\varepsilon = f - \nu_\varepsilon u^\varepsilon, \qquad \mathcal{L}^\varepsilon v^\varepsilon = f.$$

Then it follows that $u^\varepsilon - v^\varepsilon = -\mathcal{G}_\varepsilon \nu_\varepsilon u^\varepsilon = -\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon - \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon)$. Iterate this relation another time; we get the truncated Neumann series

$$u^\varepsilon - v^\varepsilon = -\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon + \mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon + \mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon). \tag{4-11}$$

In particular, the fluctuations in $u^\varepsilon - v^\varepsilon$ can be written as

$$u^\varepsilon - \mathbb{E}u^\varepsilon = -\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon + \big(\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon - \mathbb{E}\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon\big) + \big(\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon) - \mathbb{E}\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon)\big).$$

The first term above is exactly w^ε, which has mean zero and its energy was estimated in Lemma 4.1. The next lemma provides an estimate for the energy of the second term in the above expansion.

Lemma 4.2. *Suppose that the assumptions of Theorem 2.3 are satisfied. Then there exists a constant* $C > 0$, *depending only on the universal parameters and* f, *such that*

$$\mathbb{E}\big\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon - \mathbb{E}\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon\big\|^2_{L^2(D)} \le \begin{cases} C\varepsilon^{2d} & \text{if } d = 2, 3, \\ C\varepsilon^8 |\log \varepsilon|^2 & \text{if } d = 4, \\ C\varepsilon^8 & \text{if } 5 \le d \le 7. \end{cases} \tag{4-12}$$

Let I_2^ε denote the left-hand side of (4-12); it has the expression

$$\begin{aligned} I_2^\varepsilon &= \mathbb{E}\int_D \Big(\int_{D^2} G_\varepsilon(x,y) G_\varepsilon(y,z)\big(\nu^\varepsilon(y)\nu^\varepsilon(z) - \mathbb{E}\nu^\varepsilon(y)\nu^\varepsilon(z)\big) v^\varepsilon(z)\, dz\, dy\Big)^2 dx \\ &= \int_{D^5} G_\varepsilon(x,y) G_\varepsilon(x,y') G_\varepsilon(y,z) G_\varepsilon(y',z') v^\varepsilon(z) v^\varepsilon(z') \\ &\qquad \Big(\mathbb{E}\Big(\nu\Big(\frac{y}{\varepsilon}\Big)\nu\Big(\frac{y'}{\varepsilon}\Big)\nu\Big(\frac{z}{\varepsilon}\Big)\nu\Big(\frac{z'}{\varepsilon}\Big)\Big) - R\Big(\frac{y-z}{\varepsilon}\Big) R\Big(\frac{y'-z'}{\varepsilon}\Big)\Big)\, dz'\, dy'\, dz\, dy\, dx. \end{aligned}$$

It is then evident that we need to estimate certain fourth-order moments of $\nu(x,\omega)$, namely, the function

$$\Psi_\nu(x,y,t,s) := \mathbb{E}\nu(x)\nu(y)\nu(t)\nu(s) - \big(\mathbb{E}\nu(x)\nu(y)\big)\big(\mathbb{E}\nu(t)\nu(s)\big). \tag{4-13}$$

Were ν a Gaussian random field, its fourth-order moments would decompose as a sum of products of pairs of R. This nice property does not hold for general random fields; however, the following estimate for ρ-mixing random fields provides almost the same convenience.

Lemma 4.3. *Suppose* $\nu(x,\omega)$ *is a random field with maximal correlation function* ϱ *defined as in* (2-6). *Then*

$$|\Psi_\nu(x,t,y,s)| \le \vartheta(|x-t|)\vartheta(|y-s|) + \vartheta(|x-s|)\vartheta(|y-t|), \tag{4-14}$$

where $\vartheta(r) = (K\varrho(r/3))^{1/2}$, *with* $K = 4\|\nu\|_{L^\infty(\Omega\times D)}$.

We refer to [Hairer et al. 2013] for the proof of this lemma. Estimates of this type based on mixing properties already appeared in [Bal 2008]. We refer to [Bal and Jing 2011] for an alternative way to control terms like Ψ_ν, and to Section 7 for some comments on the connection of condition (S) with the lemma above.

Proof of Lemma 4.2. Integrate over x in the expression of I_2^ε, and apply the estimates (3-11), (4-7) and (4-14). We find, for $d \geq 3$,

$$I_2^\varepsilon \leq C\Biggl(\int_{D^4} \frac{\bigl(1+\mathbb{1}_{d=4}\bigl|\log|y-y'|\bigr|\bigr)|v^\varepsilon(z)v^\varepsilon(z')|}{|y-y'|^{(d-4)\vee 0}|y-z|^{d-2}|y'-z'|^{d-2}}\,\vartheta\Bigl(\frac{y-y'}{\varepsilon}\Bigr)\vartheta\Bigl(\frac{z-z'}{\varepsilon}\Bigr)\,dz'\,dy'\,dz\,dy + \int_{D^4} \frac{\bigl(1+\mathbb{1}_{d=4}\bigl|\log|y-y'|\bigr|\bigr)|v^\varepsilon(z)v^\varepsilon(z')|}{|y-y'|^{(d-4)\vee 0}|y-z|^{d-2}|y'-z'|^{d-2}}\,\vartheta\Bigl(\frac{y-z'}{\varepsilon}\Bigr)\vartheta\Bigl(\frac{z-y'}{\varepsilon}\Bigr)\,dz'\,dy'\,dz\,dy\Biggr).$$

For $d=2$, the terms $|y-z|^{-(d-2)}$ and $|y'-z'|^{-(d-2)}$ above should be replaced by $1+\bigl|\log|y-z|\bigr|$ and $1+\bigl|\log|y'-z'|\bigr|$ respectively. Let I_{21}^ε and I_{22}^ε denote the two terms on the right-hand side of the estimate above. In the following, we set ρ to be the diameter of D.

Estimate of I_{21}^ε. We use the change of variables

$$\frac{y-y'}{\varepsilon}\mapsto y,\quad \frac{z-z'}{\varepsilon}\mapsto z,\quad y'-z'\mapsto y',\quad z'\mapsto z'.$$

Then the integral in I_{21}^ε becomes, for $d \geq 3$,

$$\frac{C\varepsilon^{2d}}{\varepsilon^{(d-4)\vee 0}}\int_{B_{\rho/\varepsilon}^2} dy\,dz\int_{B_\rho} dy'\int_D dz'\,\frac{\bigl(1+\mathbb{1}_{d=4}\bigl|\log|\varepsilon y|\bigr|\bigr)|v^\varepsilon(z')v^\varepsilon(z'+\varepsilon z)|}{|y|^{(d-4)\vee 0}|y'+\varepsilon(y-z)|^{d-2}|y'|^{d-2}}\vartheta(y)\vartheta(z).$$

We integrate over y' first and apply (4-6), then integrate over z' and obtain

$$I_{21}^\varepsilon \leq C\|v^\varepsilon\|_{L^2}^2\varepsilon^{2d-2(d-4)\vee 0}\int_{B_{\rho/\varepsilon}^2}\frac{\bigl(1+\mathbb{1}_{d=4}\bigl|\log|\varepsilon y|\bigr|\bigr)\bigl(1+\mathbb{1}_{d=4}\bigl|\log|\varepsilon(y-z)|\bigr|\bigr)\vartheta(y)\vartheta(z)}{|y|^{(d-4)\vee 0}|y-z|^{(d-4)\vee 0}}\,dy\,dz.$$

When $d=3$, the integral above is bounded because $\vartheta\in L^1(\mathbb{R}^d)$ thanks to assumption (S), and we have $I_{21}^\varepsilon \leq C\varepsilon^{2d}$. When $d=2$, the situation is similar; after the integral over y', there is again no singularity in the denominator. Hence, $I_{21}^\varepsilon \leq C\varepsilon^{2d}$.

When $d\geq 5$, by the Hardy–Littlewood–Sobolev inequality [Lieb and Loss 2001, Theorem 4.3], we have, for $p, r\in(1,\infty)$,

$$\int_{\mathbb{R}^{2d}}\frac{(\vartheta(y)/|y|^{d-4})\vartheta(z)}{|y-z|^{d-4}}\,dy\,dz \leq C\biggl\|\frac{\vartheta(y)}{|y|^{d-4}}\biggr\|_{L^p(\mathbb{R}^d)}\|\vartheta\|_{L^r(\mathbb{R}^d)},\qquad \frac{1}{p}+\frac{d-4}{d}+\frac{1}{r}=2.$$

Take $p=d/(4+\delta)$ and $r=d/(d-\delta)$ for any $(d-8)\vee 0<\delta<d-4$. Then because $\vartheta\in L^\infty\cap L^1(\mathbb{R}^d)$ and $|y|^{4-d}\in L^p(B_1)$, the above is finite and we have $I_{21}^\varepsilon\leq C\varepsilon^8$.

When $d=4$, we need to control the integral

$$\int_{B_{\rho/\varepsilon}^2}\bigl(1+\bigl|\log|\varepsilon y|\bigr|\bigr)\bigl(1+\bigl|\log|\varepsilon(y-z)|\bigr|\bigr)\vartheta(y)\vartheta(z)\,dy\,dz,$$

where $D^* = \{y - y' - z + z' : y, y', z, z' \in D\}$ is some bounded region formed by certain combinations of points in D. As a result, the logarithmic terms are bounded away from the poles. Hence, the above integral is bounded by $O(|\log\varepsilon|^2)$, and $I^\varepsilon_{21} \le C\varepsilon^8|\log\varepsilon|^2$.

Estimate of I^ε_{22}. We apply the change of variables

$$\frac{y-z'}{\varepsilon} \mapsto y, \quad \frac{y'-z}{\varepsilon} \mapsto y', \quad z - z' \mapsto z, \quad z' \mapsto z'.$$

Then the integral in I^ε_{22} becomes, for $d = 2$,

$$C\varepsilon^{2d}\int_{B^2_{\rho/\varepsilon}} dy\,dy'\int_D dz'\int_{B_\rho} dz\big|v^\varepsilon(z')v^\varepsilon(z'+z)\big|\big(1+\big|\log|z-\varepsilon y|\big|\big)(1+\log|z-\varepsilon y'|)\vartheta(y)\vartheta(y').$$

Integrate over z', z and then over y' and y. We find that $I^\varepsilon_{22} \le C\varepsilon^{2d}$. For $d \ge 3$, the same change of variables transforms I^ε_{22} to

$$C\varepsilon^{2d}\int_{B^2_{\rho/\varepsilon}} dy\,dy'\int_{B_\rho} dz\int_D dz'\frac{\big(1+\mathbb{1}_{d=4}\big|\log|z-\varepsilon(y-y')|\big|\big)|v^\varepsilon(z')v^\varepsilon(z'+z)|}{|z-\varepsilon(y-y')|^{(d-4)\vee 0}|z-\varepsilon y|^{d-2}|z+\varepsilon y'|^{d-2}}\vartheta(y)\vartheta(y').$$

After an integration over z', we only need to control

$$C\varepsilon^{2d}\int_{B^2_{\rho/\varepsilon}} dy\,dy'\int_{B_\rho} dz\frac{\big(1+\mathbb{1}_{d=4}\big|\log|z-\varepsilon(y-y')|\big|\big)}{|z-\varepsilon(y-y')|^{(d-4)\vee 0}|z-\varepsilon y|^{d-2}|z+\varepsilon y'|^{d-2}}\vartheta(y)\vartheta(y').$$

When $d = 3$, an integration over z removes the singularities in the denominator. Then integrating over y and y' yields that $I^\varepsilon_{22} \le C\varepsilon^{2d}$.

When $d \ge 5$, we need to control the integral; after another change of variables, $\varepsilon^{-1}z - (y - y') \mapsto z$ and $-y \mapsto y$, we have

$$\int_{\mathbb{R}^{3d}} dy\,dy'\,dz\frac{C\varepsilon^8\vartheta(y)\vartheta(y')}{|z|^{d-4}|z-y|^{d-2}|z-y'|^{d-2}} = \int_{\mathbb{R}^d} dz\frac{C\varepsilon^8|K(z)|^2}{|z|^{d-4}},$$

with $K(z) = (|y|^{-(d-2)} * \vartheta)(z)$. Since $\vartheta \in L^1 \cap L^\infty(\mathbb{R}^d)$, we have

$$|K(z)| = \int_{B_1(z)}\frac{\vartheta(y)\,dy}{|z-y|^{d-2}} + \int_{\mathbb{R}^d\setminus B_1(z)}\frac{\vartheta(y)\,dy}{|z-y|^{d-2}} \le \int_{B_1(z)}\frac{\|\vartheta\|_{L^\infty}\,dy}{|y-z|^{d-2}} + \int_{\mathbb{R}^d\setminus B_1(z)}\vartheta(y)\,dy \le C.$$

Moreover, by the Hardy–Littlewood–Sobolev inequality, we have that

$$\|K\|_{L^2(\mathbb{R}^d)} = \big\||y|^{-(d-2)} * \vartheta(y)\big\|_{L^2(\mathbb{R}^d)} \le C\|\vartheta\|_{L^{2d/(d+4)}(\mathbb{R}^d)} \le C\|\vartheta\|^{\frac{d-4}{2d}}_{L^\infty}\|\vartheta\|^{\frac{d+4}{2d}}_{L^1}.$$

Now we show that $K \in L^\infty \cap L^2(\mathbb{R}^d)$. It follows that the integral to be controlled is finite and we have $I^\varepsilon_{22} \le C\varepsilon^8$.

When $d = 4$, after the same change of variables as in the case of $d \ge 5$, we are left to control

$$\varepsilon^8\int_{B^2_{\rho/\varepsilon}} dy\,dy'\int_{B_{3\rho/\varepsilon}} dz\frac{\big(1+\big|\log|\varepsilon z|\big|\big)\vartheta(y)\vartheta(y')\,dy\,dy'\,dz}{|z-y|^2|z-y'|^2} = \varepsilon^8\int_{B_{3\rho/\varepsilon}}\big(1+\big|\log|\varepsilon z|\big|\big)(K(z))^2\,dz,$$

where $K(z) = (\mathbb{1}_{B_{\rho/\varepsilon}}(y)|y|^{-2} * \vartheta)(z)$. We verify again that $K \in L^\infty \cap L^{2-\delta}(\mathbb{R}^d)$ for any $\delta \in (0,1)$. Estimate the integral again by breaking it into pieces inside and outside B_1; we find $I_{22}^\varepsilon \le C\varepsilon^8|\log\varepsilon|$.

Combining these estimates above, we have proved (4-12). □

Moving on to the last term in the series (4-11), we observe that it cannot be controlled in the same manner as above. Indeed, the term $u^\varepsilon - v^\varepsilon$ is random and depends on $\nu(x,\omega)$ in a nonlinear way. As a result, when we move the expectation into the integral representation, like in step (4-5), we cannot get a simple closed form in terms of R.

We hence choose not to address the interaction between $u^\varepsilon - v^\varepsilon$ and the random fluctuation ν^ε in the potential directly. Instead, by an application of Minkowski's inequality, we have

$$\left\|\mathbb{E}\,\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u)\right\|_{L^2(D)} \le \mathbb{E}\left\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u)\right\|_{L^2(D)}.$$

Thus, we use the trivial bound on the $L^1(\Omega, L^2(D))$-norm of the fluctuations in $\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u)$:

$$r_2^\varepsilon := \mathbb{E}\left\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u) - \left(\mathbb{E}\,\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u)\right)\right\|_{L^2(D)} \le 2\,\mathbb{E}\left\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-u)\right\|_{L^2(D)},$$

and only control the energy of the last term in (4-11) itself, as contrast to its variance.

Lemma 4.4. *Suppose that the assumptions of Theorem 2.3 are satisfied. Then there exists some constant C, depending only on the universal parameters and f, such that*

$$\mathbb{E}\left\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\nu^\varepsilon(u^\varepsilon-v^\varepsilon)\right\|_{L^2(D)} \le \begin{cases} C\varepsilon^d & \text{if } d=2,3,\\ C\varepsilon^4|\log\varepsilon|^{\frac{3}{2}} & \text{if } d=4,\\ C\varepsilon^{6-\frac{d}{2}} & \text{if } d\ge 5.\end{cases} \tag{4-15}$$

To prove this result, we estimate the operator norm of $\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon$, which is random since ν^ε depends on ω, and combine it with the control of $u^\varepsilon - v^\varepsilon$, which was obtained earlier.

Lemma 4.5 (mean value of the operator norm $\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\|_{L^2\to L^2}$). *Under the same assumptions of Theorem 2.3, there exists some universal constant C such that*

$$\mathbb{E}\,\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon\|^2_{L^2\to L^2} \le \begin{cases} C\varepsilon^d & \text{if } d=2,3,\\ C\varepsilon^4|\log\varepsilon|^2 & \text{if } d=4,\\ C\varepsilon^{8-d} & \text{if } d\ge 5.\end{cases} \tag{4-16}$$

Proof. For any $h \in L^2(D)$, we have

$$\|\mathcal{G}_\varepsilon\nu^\varepsilon\mathcal{G}_\varepsilon h\|^2_{L^2} = \int_D\left(\int_{D^2} G_\varepsilon(x,y)\nu^\varepsilon(y)G_\varepsilon(y,z)h(z)\,dz\,dy\right)^2 dx.$$

Note that for almost every fixed $x \in D$,

$$\left|\int_{D^2} G_\varepsilon(x,y)\nu^\varepsilon(y)G_\varepsilon(y,z)h(z)\,dz\,dy\right| \le \|h\|_{L^2}\left\|\int_D G_\varepsilon(x,y)\nu^\varepsilon(y)G_\varepsilon(y,\cdot)\,dy\right\|_{L^2}.$$

It then follows that

$$\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon\|^2_{L^2\to L^2}(\omega) \le \int_{D^2}\biggl(\int_D G_\varepsilon(x,y)\nu^\varepsilon(y,\omega)G_\varepsilon(y,z)\,dy\biggr)^2\,dz\,dx.$$

Taking the expectation, we find

$$\mathbb{E}\,\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon\|^2_{L^2\to L^2} \le \int_{D^4} G_\varepsilon(x,y)G_\varepsilon(x,\eta)R\Bigl(\frac{y-\eta}{\varepsilon}\Bigr)G_\varepsilon(y,z)G_\varepsilon(\eta,z)\,dy\,d\eta\,dz\,dx.$$

Integrate over z- and x-variables first. Using (3-11) and (4-6), we find that the integrals over x- and z-variables are estimated as in (4-7). Then we have

$$\mathbb{E}\,\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon\|^2_{L^2\to L^2} \le C\int_{D^2}\biggl(\frac{1+\mathbb{1}_{d=4}\bigl|\log|y-\eta|\bigr|}{|y-\eta|^{(d-4)\vee 0}}\biggr)^2\,\Bigl|R\Bigl(\frac{y-\eta}{\varepsilon}\Bigr)\Bigr|\,dy\,d\eta.$$

Change variables in the above integral and carry out the analysis as before. We find that (4-16) holds. Note that the estimates become useless for $d \ge 8$. □

Proof of Lemma 4.4. For each $\omega \in \Omega$, we have

$$\bigl\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon(u^\varepsilon - v^\varepsilon)\bigr\|_{L^2} \le M\,\|\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon\|_{L^2\to L^2}\,\|u^\varepsilon - v^\varepsilon\|_{L^2},$$

where M is the uniform bound on the random potential in (2-4). Take the expectation and then the desired estimate follows from (4-16), (4-2) and (4-3). □

4B. ***Scaling factor of the random fluctuations in the weak topology.*** In this section we aim to find the correct scaling factor such that the random fluctuation $u^\varepsilon - \mathbb{E}\,u^\varepsilon$, normalized properly according to this factor, converges with respect to the weak topology. For that purpose, we fix an arbitrary $\varphi \in L^2(D)$ with unit norm, and estimate $\mathbb{E}(u^\varepsilon - \mathbb{E}\,u^\varepsilon, \varphi)^2$.

Using the series expansion formula (4-11), we have

$$\begin{aligned}(u^\varepsilon - \mathbb{E}\,u^\varepsilon,\varphi) = -(\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon,\varphi) &+ \bigl(\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon - \mathbb{E}(\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon),\varphi\bigr)\\ &+ \bigl(\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon(u^\varepsilon - v^\varepsilon) - \mathbb{E}(\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon(u^\varepsilon - v^\varepsilon)),\varphi\bigr).\end{aligned}$$

Since the operators $\mathcal{G}_\varepsilon$ and $\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon$ are self-adjoint on $L^2(D)$, we can move them to φ. Set $\psi^\varepsilon = \mathcal{G}_\varepsilon\varphi$. The above expression becomes

$$\begin{aligned}(u^\varepsilon - \mathbb{E}u^\varepsilon,\varphi) &= -(\nu^\varepsilon v^\varepsilon,\psi^\varepsilon) + \bigl((\nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon,\psi^\varepsilon) - \mathbb{E}(\nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon,\psi^\varepsilon)\bigr)\\ &\qquad + \bigl((\nu^\varepsilon(u^\varepsilon - v^\varepsilon),\mathcal{G}_\varepsilon \nu^\varepsilon\psi^\varepsilon) - \mathbb{E}(\nu^\varepsilon(u^\varepsilon - v^\varepsilon),\mathcal{G}_\varepsilon \nu^\varepsilon\psi^\varepsilon)\bigr)\\ &:= I_1^\varepsilon + (I_2^\varepsilon - \mathbb{E}\,I_2^\varepsilon) + (I_3^\varepsilon - \mathbb{E}\,I_3^\varepsilon).\end{aligned} \tag{4-17}$$

The aim now is to control the variances of I_j^ε, with $j = 1, 2, 3$.

Estimate for I_1^ε. For I_1^ε, which is mean-zero, we have

$$\begin{aligned}
\mathbb{E}(I_1^\varepsilon)^2 &= \mathbb{E}\Big(\int_D v^\varepsilon(x)v^\varepsilon(x)\psi^\varepsilon(x)\,dx\Big)^2 = \int_{D^2} R\Big(\frac{x-y}{\varepsilon}\Big)v^\varepsilon(x)v^\varepsilon(y)\psi^\varepsilon(x)\psi^\varepsilon(y)\,dx\,dy \\
&= \int_{\mathbb{R}^d}\big[R^\varepsilon * \big(v^\varepsilon(y)\psi^\varepsilon(y)\mathbb{1}_D(y)\big)\big](x)\,v^\varepsilon(x)\psi^\varepsilon(x)\mathbb{1}_D(x)\,dx \\
&\le C\|R^\varepsilon\|_{L^1(\mathbb{R}^d)}\|v^\varepsilon\psi^\varepsilon\mathbb{1}_D\|^2_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

Here, $R^\varepsilon(y) = R\big(\frac{y}{\varepsilon}\big)$ is a shorthand notation. To obtain the last inequality, we applied Hölder's and Young's inequalities. Note that $\|R^\varepsilon\|_{L^1(\mathbb{R}^d)} = \varepsilon^d\|R\|_{L^1(\mathbb{R}^d)}$. Note also that $f, \varphi \in L^2(D)$ implies that $v^\varepsilon, \psi^\varepsilon \in H^2(D)$, which is embedded in $L^4(D)$ for all $2 \le d \le 7$. As a result, we conclude that $\mathbb{E}\,|I_1^\varepsilon| \le C\varepsilon^{d/2}$.

Estimate for $\mathrm{Var}(I_2^\varepsilon)$. Before calculating the variance of I_2^ε, we first check that $\|I_2^\varepsilon\|_{L^2(\Omega)}$ can have size larger than $\varepsilon^{d/2}$ for $d \ge 4$. By direct computation, for $d \ge 3$,

$$\begin{aligned}
\mathbb{E}(I_2^\varepsilon)^2 &= \int_{D^4} R\Big(\frac{x-y}{\varepsilon}\Big)R\Big(\frac{x'-y'}{\varepsilon}\Big)G_\varepsilon(x,y)G_\varepsilon(x',y')v^\varepsilon(y)v^\varepsilon(y')\psi^\varepsilon(x)\psi^\varepsilon(x')\,dx'\,dy'\,dx\,dy \\
&\lesssim \int_{D^4}\Big|R\Big(\frac{x-y}{\varepsilon}\Big)R\Big(\frac{x'-y'}{\varepsilon}\Big)\Big|\frac{|v^\varepsilon(y)v^\varepsilon(y')\psi^\varepsilon(x)\psi^\varepsilon(x')|}{|x-y|^{d-2}|x'-y'|^{d-2}}\,dy'\,dx'\,dy\,dx.
\end{aligned}\tag{4-18}$$

For $d = 2$, the last integral above should be replaced by

$$\int_{D^4}\Big|R\Big(\frac{x-y}{\varepsilon}\Big)R\Big(\frac{x'-y'}{\varepsilon}\Big)v^\varepsilon(y)v^\varepsilon(y')\psi^\varepsilon(x)\psi^\varepsilon(x')\big(1+\big|\log|x-y|\big|\big)\big(1+\big|\log|x'-y'|\big|\big)\Big|\,dy'\,dx'\,dy\,dx.$$

After the change of variables

$$\frac{x-y}{\varepsilon}\mapsto x, \quad \frac{x'-y'}{\varepsilon}\mapsto x', \quad y\to y, \quad y'\to y',$$

the integral to be controlled, for $d \ge 3$, becomes

$$\varepsilon^4\int_{B^2_{\rho/\varepsilon}}\int_{D^2}|R(x)R(x')|\frac{\big|v^\varepsilon(y)\psi^\varepsilon(y+\varepsilon x)v^\varepsilon(y')\psi^\varepsilon(y'+\varepsilon x')\big|}{|x|^{d-2}|x'|^{d-2}}\,dy'\,dy\,dx'\,dx.$$

Integrating over y and y' and then over x and x', we find that the integral above is finite. Hence, $\mathbb{E}(I_2^\varepsilon)^2$ is of order ε^4 when $d \ge 3$. When $d = 2$, the change of variables in the logarithmic functions yields the term $|\log\varepsilon|^2$, and we have $\mathbb{E}(I_2^\varepsilon)^2$ is of order $\varepsilon^4|\log\varepsilon|^2$. This shows that the second term in (4-11), i.e., I_2^ε, is larger than or comparable to $\varepsilon^{d/2}$ for $d \ge 4$.

We show next that the variance of I_2^ε, however, is smaller than $\varepsilon^{d/2}$ in all dimensions. Using the definition of Ψ_ν in (4-13) and the estimate in Lemma 4.3, we bound $\mathbb{E}(I_2^\varepsilon - \mathbb{E} I_2^\varepsilon)^2 = \mathrm{Var}(I_2^\varepsilon)$, for $d \ge 3$, by

$$\begin{aligned}\mathrm{Var}(I_2^\varepsilon) &= \int_{D^4} \Psi_\nu\Big(\frac{x}{\varepsilon}, \frac{y}{\varepsilon}, \frac{x'}{\varepsilon}, \frac{y'}{\varepsilon}\Big) G_\varepsilon(x,y) G_\varepsilon(x',y') v^\varepsilon(y) v^\varepsilon(y') \psi^\varepsilon(x) \psi^\varepsilon(x')\, dx'\, dy'\, dx\, dy \\ &\le C \int_{D^4} \vartheta\Big(\frac{x-x'}{\varepsilon}\Big) \vartheta\Big(\frac{y-y'}{\varepsilon}\Big) \frac{\big|v^\varepsilon(y) v^\varepsilon(y') \psi^\varepsilon(x) \psi^\varepsilon(x')\big|}{|x-y|^{d-2} |x'-y'|^{d-2}}\, dx'\, dy'\, dx\, dy \\ &\qquad + C \int_{D^4} \vartheta\Big(\frac{x-y'}{\varepsilon}\Big) \vartheta\Big(\frac{y-x'}{\varepsilon}\Big) \frac{\big|v^\varepsilon(y) v^\varepsilon(y') \psi^\varepsilon(x) \psi^\varepsilon(x')\big|}{|x-y|^{d-2} |x'-y'|^{d-2}}\, dx'\, dy'\, dx\, dy.\end{aligned}$$

The second integral above is essentially the same with the first one if we interchange x' and y'. Hence, we focus only on the first one. After the change of variables

$$\frac{x-x'}{\varepsilon} \mapsto x, \quad \frac{y-y'}{\varepsilon} \mapsto y, \quad y - x' \mapsto x', \quad y' \mapsto y',$$

the first integral becomes

$$C\varepsilon^{2d} \int_{\mathbb{R}^{2d}} dx\, dy \int_{B_\rho} dx' \int_D dy' \vartheta(x) \vartheta(y) \frac{\big|v^\varepsilon(y'+\varepsilon y) v^\varepsilon(y') \psi^\varepsilon(y'-x'+\varepsilon x+\varepsilon y) \psi^\varepsilon(y'-x'+\varepsilon y)\big|}{|x'-\varepsilon x|^{d-2} |x'-\varepsilon y|^{d-2}}.$$

Integrate over y' first and use the fact that $\|v^\varepsilon\|_{L^4} \le C$ and $\|\psi^\varepsilon\|_{L^4(D)} \le C$. Then the above integral is bounded by

$$\int_{\mathbb{R}^{2d}} dx\, dy \int_{B_\rho} \vartheta(x) \vartheta(y) \frac{C\varepsilon^{2d}\, dx'}{|x'-\varepsilon x|^{d-2} |x'-\varepsilon y|^{d-2}} \le \int_{\mathbb{R}^{3d}} \frac{C\varepsilon^{2d-(d-4)\vee 0} \vartheta(x) \vartheta(y)\, dx'\, dx\, dy}{|x-y|^{(d-4)\vee 0}}$$

for $d \ne 4$, where we integrated over x' and used (4-6) to have the inequality. The resulting integral is clearly finite. Hence we conclude that $\mathrm{Var}(I_2^\varepsilon) \le C\varepsilon^{2d}$ for $d = 3$ and that it is of order ε^{d+4} for $d \ge 5$.

When $d = 2$, there is only logarithmic singularity to start with in the expression of $\mathrm{Var}(I_2^\varepsilon)$, and we find $\mathrm{Var}(I_2^\varepsilon) \le C\varepsilon^{2d}$.

When $d = 4$, the integral that remains after we integrate over x' has a term of the form

$$\big(\log|\varepsilon(x-y)|\big) \mathbb{1}_{\varepsilon(x-y) \in B_{2\rho}}.$$

It follows then that $\mathrm{Var}(I_2^\varepsilon) \le C\varepsilon^8 |\log \varepsilon|$.

To summarize, for $d \ge 2$, we have $\mathbb{E}|I_2^\varepsilon - \mathbb{E}\, I_2^\varepsilon| \ll \mathbb{E}|I_1^\varepsilon|$. That is, when the series expansion is integrated against test functions and the mean is removed, the second term is much smaller than the leading term.

Estimate for I_3^ε. For the last term, we control it by the crude estimate $\mathbb{E}(I_3^\varepsilon - \mathbb{E}\, I_3^\varepsilon)^2 \le 2\mathbb{E}(I_3^\varepsilon)^2$. From the expression $I_3^\varepsilon = (v^\varepsilon(u^\varepsilon - v^\varepsilon), \mathcal{G}_\varepsilon v^\varepsilon \psi^\varepsilon)$, we have

$$\mathbb{E}|I_3^\varepsilon| \le \mathbb{E}\big(\|v^\varepsilon\|_{L^\infty} \|u^\varepsilon - v^\varepsilon\|_{L^2} \|\mathcal{G}_\varepsilon v^\varepsilon \psi^\varepsilon\|_{L^2}\big) \le C\big(\mathbb{E}\|u^\varepsilon - v^\varepsilon\|_{L^2}^2\, \mathbb{E}\|\mathcal{G}_\varepsilon v^\varepsilon \psi^\varepsilon\|_{L^2}^2\big)^{\frac{1}{2}}$$

since $\mathcal{G}_\varepsilon v^\varepsilon \psi^\varepsilon$ is exactly of the form of w^ε defined in (4-1). Owing to (4-2) and Lemma 4.1, we conclude that $\mathbb{E}\,|I_3^\varepsilon|$ is of order ε^d for $d = 2, 3$, of order $\varepsilon^4 |\log \varepsilon|$ for $d = 4$, and of order ε^4 for $d \ge 5$. Hence, for all $2 \le d \le 7$, the truncation term in the Neumann series, with respect to the weak topology, has a scaling factor that is smaller than that of the leading term (which is of order $\varepsilon^{d/2}$).

Remark 4.6. We find that for $2 \le d \le 7$, the random fluctuation $u^\varepsilon - \mathbb{E}\, u^\varepsilon$ scales like $\varepsilon^{d/2}$ when integrated against test functions, and the leading term is the dominating one. We do not expect the dimension constraint $d \le 7$ to be intrinsic. Firstly, it is related to the fact that we stopped at the second-order iteration in the Neumann series, and had to control the last term by the crude estimate given by the Minkowski inequality (not taking advantage of removing the mean). Secondly, it is also needed when we claim that $\psi^\varepsilon = \mathcal{G}_\varepsilon \nu^\varepsilon \varphi$ is in $L^4(D)$. In general, if we assume a stronger condition, namely $f \in C(\bar{D})$, then v^ε is always bounded, and we only need $\psi^\varepsilon \in L^2(D)$, which holds in all dimensions if $\varphi \in L^2(D)$.

We conclude this section by collecting the facts obtained above to give a proof of Theorem 2.3.

Proof of Theorem 2.3. Let v^ε be as defined in (3-7). In view of (3-10) and the Minkowski inequality, we have

$$\mathbb{E}\, \|u^\varepsilon - \mathbb{E}\, u^\varepsilon\|_{L^2}^2 \le \mathbb{E}\big(2\|u^\varepsilon - v^\varepsilon\|_{L^2}^2 + 2\,\|\mathbb{E}(u^\varepsilon - v^\varepsilon)\|_{L^2}^2\big) \le 4\,\mathbb{E}\,\|u^\varepsilon - v^\varepsilon\|_{L^2}^2.$$

Owing to (4-2) and Lemma 4.1, we have (2-13).

In view of (3-8), (4-2), Lemma 4.1 and Theorem 3.3, we have

$$\mathbb{E}\,\|u^\varepsilon - u\|_{L^2} \le \mathbb{E}\,\|u^\varepsilon - v^\varepsilon\|_{L^2} + \|v^\varepsilon - u\|_{L^2} \le C\varepsilon \|f\|_{L^2}.$$

This proves (2-12).

Finally, to estimate $\mathbb{E}|(u^\varepsilon - \mathbb{E}u^\varepsilon, \varphi)_{L^2}|$ for an arbitrary field $\varphi \in L^2(D)$, without loss of generality we can assume $\|\varphi\|_{L^2} = 1$. Then this term is precisely what was studied immediately above. With I_j^ε, where $j = 1, 2, 3$, defined earlier, we have showed that for $2 \le d \le 7$, we have $\mathbb{E}\,|\sum_{j=1}^3 (I_j^\varepsilon - \mathbb{E}\, I_j^\varepsilon)| \le C\varepsilon^{d/2}$, which is precisely (2-14). □

5. Limiting distribution of the random fluctuation

In this section, we study the limiting distribution of the scaled random fluctuation, $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}\, u^\varepsilon)$, in functional spaces. As mentioned earlier, the choice of space depends on dimension. When $d = 1$, convergence in law in $C(D)$ of the random fluctuation was proved in [Bourgeat and Piatnitski 1999; Bal 2008]. We prove Theorem 2.4 below, which establishes convergence in law of the random fluctuation in $L^2(D)$ for $d = 2, 3$ and in $H^{-1}(D)$ for $d = 4, 5$.

Multiplying $\varepsilon^{-d/2}$ to the series expansion (4-11), we obtain the following expression for the scaled random fluctuation:

$$-\frac{\mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon}{\sqrt{\varepsilon^d}} + \frac{\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon - \mathbb{E}\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon v^\varepsilon}{\sqrt{\varepsilon^d}} + \frac{\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon) - \mathbb{E}\mathcal{G}_\varepsilon \nu_\varepsilon \mathcal{G}_\varepsilon \nu_\varepsilon (u^\varepsilon - v^\varepsilon)}{\sqrt{\varepsilon^d}}. \tag{5-1}$$

Our strategy, as in [Bal 2008; Bal and Jing 2011], is to prove that the leading term $X^\varepsilon = -\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon$ contributes and converges in law to the right distribution depicted by Theorem 2.4, and show that the other terms converge in stronger mode to the zero function and hence have no contribution to the limiting law.

At the purely formal level, all these steps are the same as in the setting of nonoscillatory diffusion coefficients. Indeed, we already established controls for the second and last terms above in the previous section. Moreover, the ε-dependence in $\mathcal{G}_\varepsilon$ and v^ε is not a problem, as we will see later, for the convergence of the characteristic functions of X^ε, thanks to the fact that $\mathcal{G}_\varepsilon\varphi \to \mathcal{G}\varphi$ in L^2 for any $\varphi \in H^{-1}(D)$. This dependence, however, does impose difficulty on showing the tightness of the measures of $\{X^\varepsilon\}_\varepsilon$. As discussed in Section 7, the old approach for tightness in [Bal et al. 2012] fails and new ideas are needed.

Our new approach is to use some nonoptimal but convenient tightness criteria, described in Theorems A.1 and A.2, for probability measures on $H^k(D)$ that are induced by processes in $H^{k+s}(D)$, with $k = -1, 0$ and $s > 0$. Since we do need s to be fractional in $(0, 1)$, we recall some definitions regarding fractional Sobolev spaces; see [Di Nezza et al. 2012] for reference. Given an open set $K \subset \mathbb{R}^d$, the fractional Sobolev space $H^s_0(K)$, for $s \in (0, 1)$, is the closure of $C^\infty_0(K)$ in the norm

$$\|u\|^2_{H^s(K)} := \|u\|^2_{L^2(K)} + \int_{K^2} \frac{|u(x)-u(y)|^2}{|x-y|^{d+2s}}\, dx\, dy.$$

When $K = \mathbb{R}^d$, an equivalent norm for $u \in H^s(\mathbb{R}^d)$ is

$$\|u\|^2_{H^s(\mathbb{R}^d)} := \int_{\mathbb{R}^d} (1+|\xi|^2)^s\, |\mathcal{F}u|^2(\xi)\, d\xi. \tag{5-2}$$

Moreover, for $s \in (0, 1)$, the space $H^{-s}(K)$ is defined to be the dual space $(H^s_0(K))'$, and in particular when $K = \mathbb{R}^d$, the above norm for $H^{-s}(\mathbb{R}^d)$ is still valid.

5A. ***Limiting distribution in $L^2(D)$ for dimensions two and three.*** For $d = 2, 3$, we prove that the leading term X^ε in (4-11) converges in law in $L^2(D)$ and show that the other terms vanish in the limit. The next lemma, together with Theorem A.1, yields tightness of X^ε, which is the key step.

Lemma 5.1. *Suppose that the conditions of Theorem 2.3 are satisfied. Assume further that $d = 2, 3$. Then for any $s\in\big(0, \frac{1}{2}\big)$, there exists a constant C, depending only on the universal parameters and s, such that*

$$\mathbb{E}\,\|\varepsilon^{-\frac{d}{2}}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon\|^2_{H^s} \le C. \tag{5-3}$$

Proof. For each fixed $\omega \in \Omega$ and $\varepsilon > 0$, we know that $\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon$ belongs to $H^1_0(D)$ and hence also to $H^s_0(D)$ for any $s \in (0, 1)$. In particular, its H^s-seminorm has the expression

$$[\varepsilon^{-\frac{d}{2}}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon]^2_{H^s(D)} = \frac{1}{\varepsilon^d}\int_{D^2} \frac{\big|(\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon)(x) - (\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon)(y)\big|^2}{|x-y|^{d+2s}}\, dy\, dx.$$

Taking the expectation and using the L^4-bounds of v^ε, we have

$$\mathbb{E}[\varepsilon^{-\frac{d}{2}}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon]^2_{H^s(D)} \le \frac{C}{\varepsilon^d}\int_{D^4} \frac{\big|\big(G_\varepsilon(x,z)-G_\varepsilon(y,z)\big)\big(G_\varepsilon(x,\xi)-G_\varepsilon(y,\xi)\big)\big|}{|x-y|^{d+2s}} \left|R\left(\frac{z-\xi}{\varepsilon}\right)\right| d\xi\, dz\, dy\, dx.$$

We claim: there exists C, depending only on the universal parameters and s, such that for all $\xi, z \in D$,

$$\int_{D^2} \frac{\big|\big(G_\varepsilon(x,z)-G_\varepsilon(y,z)\big)\big(G_\varepsilon(x,\xi)-G_\varepsilon(y,\xi)\big)\big|}{|x-y|^{d+2s}}\, dy\, dx \le C. \tag{5-4}$$

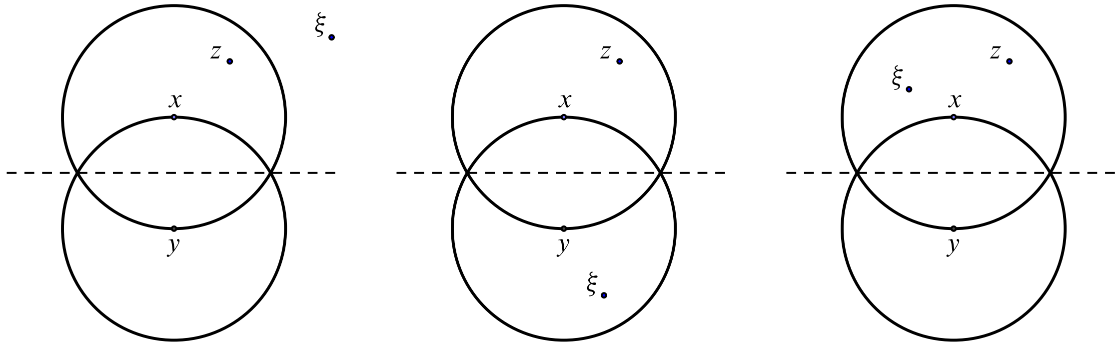

Figure 1. Decomposition criteria of the domain of integration based on the relative position between four points. Left: $(x, y) \in D_1^2$; middle: $(x, y) \in D_2^2$; right: $(x, y) \in D_3^2$.

We decompose the integration region D^2 into three parts D_j^2, with $j = 1, 2, 3$, as follows: in D_1^2, one of the points in $\{z, \xi\}$, namely ξ without loss of generality, lies outside $B_\rho(x) \cup B_\rho(y)$, where $\rho = |x - y|$; in D_2^2, one of the points, namely z without loss of generality, lies in $B_\rho(x)$ and satisfies $|z - x| \le |z - y|$ and at the same time $\xi \in B_\rho(y)$ and $|\xi - y| \le |\xi - x|$; in D_3^2, we have that ξ and z cluster around one of the points in $\{x, y\}$; without loss of generality, assume this point is x, so $z, \xi \in B_\rho(x) \cap \{\eta : |\eta - x| < |\eta - y|\}$. In Figure 1, the relative positions between $\{x, y, z, \xi\}$ are illustrated for each case.

Let I_j be the integral over D_j^2 of the integrand in (5-4). We estimate I_j, with $j = 1, 2, 3$, separately and we focus on the case of $d = 3$. It is clear that when $d = 2$, the only change is that the Green's function has logarithmic bound, and the analysis below can be adapted.

On D_1^2, without loss of generality, we assume that $|z - x| \le |z - y|$ (if otherwise, we would switch the roles of x and y). Hence $|G_\varepsilon(x, z) - G_\varepsilon(y, z)| \le C|x - z|^{2-d}$. By the mean value theorem,

$$|G_\varepsilon(x, \xi) - G_\varepsilon(y, \xi)| \le |\nabla G_\varepsilon(\eta, \xi)|\,|x - y| \quad \text{for some } \eta \text{ between } x \text{ and } y.$$

By the gradient bound (3-12) and the fact that $|\eta - \xi| \ge |y - \xi|/2$, we have

$$|\nabla G_\varepsilon(\eta, \xi)| \le \frac{C}{|\eta - \xi|^{d-1}} \le \frac{C}{|y - \xi|^{d-1}}, \quad \text{hence} \quad |G_\varepsilon(x, \xi) - G_\varepsilon(y, \xi)| \le \frac{C|x - y|}{|y - \xi|^{d-1}}.$$

As a result, we have,

$$I_1 \le \int_{D_1^2} \frac{C}{|x - y|^{d+2s-1}} \frac{1}{|x - z|^{d-2}} \frac{1}{|y - \xi|^{d-1}}\, dx\, dy.$$

Integrate over x first and then over y, using (4-6) in each step; we find that as long as $0 < s < \frac{1}{2}$, we have $I_1 \le C$ for some C that only depends on the universal parameters and s.

On D_2^2, we have $|G_\varepsilon(x, z) - G_\varepsilon(y, z)| \le C|x - z|^{2-d}$ and $|G_\varepsilon(x, \xi) - G_\varepsilon(y, \xi)| \le C|y - \xi|^{2-d}$. At the same time, $|x - z| \le |x - y|$ and $|y - \xi| \le |x - y|$, so we may split the singularity into the integrals over x and y so that each of them is essentially not singular. That is,

$$I_2 \le \int_{D_2^2} \frac{C}{|x - z|^{\frac{d}{2}+s}|y - \xi|^{\frac{d}{2}+s}} \frac{1}{|x - z|^{d-2}} \frac{1}{|y - \xi|^{d-2}}\, dx\, dy.$$

We note that the integral above can be separated, and as long as $0 < s < 2 - \frac{d}{2} = \frac{1}{2}$, each integral is finite and hence $I_2 \le C$.

On D_3^2, we assume without loss of generality that z and ξ cluster around x. Then we have

$$|G_\varepsilon(x,\eta) - G_\varepsilon(y,\eta)| \le C|x-\eta|^{2-d}$$

for $\eta \in \{z, \xi\}$. At the same time, $|x-y| > |y-z|$. As a result, we have

$$I_3 \le \int_{D_3^2} \frac{C}{|y-z|^{d-\tau}|x-z|^{2s+\tau}} \frac{1}{|x-z|^{d-2}} \frac{1}{|x-\xi|^{d-2}}\, dx\, dy.$$

We choose $\tau > 0$ so the integral over y is uniformly bounded. The integral over x is also bounded as long as $2s + \tau < (4-d) \wedge 2 = 1$, and we have $I_3 \le C$. We note that for any $s \in \left(0, \frac{1}{2}\right)$, there exists $\tau \in (0, 1-2s)$ satisfying the constraint $2s + \tau < 1$.

The above bounds are uniform in δ. Therefore, taking the limit $\delta \to 0$, we prove (5-4). Integrate over z and ξ in the integral expression of $\mathbb{E}[\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon]^2_{H^s}$; in particular, integrating $R(\cdot/\varepsilon)$ yields a factor of ε^d that cancels the one in the denominator. We conclude that $\mathbb{E}[\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon]^2_{H^s} \le C$ for each fixed $s \in \left(0, \frac{1}{2}\right)$. Combining this with $E\,\|\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon\|^2_{L^2} \le C$, which is due to (4-1) for $d = 2, 3$, we prove (5-3). □

Remark 5.2. The key step in the proof above is to derive (5-4), which concerns only the Green's function G_ε and hence is obtained from a purely deterministic argument. Indeed, the scaling factor $\varepsilon^{-d/2}$ plays a role only afterward when we integrate against R^ε, and it disappears in the final estimate because it is the right scaling for integrals of R^ε. In Section 6, where we consider the case of long-range correlated random potential $q(x,\omega)$, the scaling in X^ε will be different, but the tightness of (the measures of) X^ε, with the right scaling, is obtained in the same way as above.

Next we address the convergence of the characteristic function of the measure of X^ε. In view of Theorem 2.1, this amounts to proving this:

Lemma 5.3. *Assume* (S) *holds. For any fixed* $\varphi \in L^2(D)$, *we have*

$$-\frac{1}{\sqrt{\varepsilon^d}}(\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon, \varphi)_{L^2} \xrightarrow{\text{distribution}} \mathcal{N}(0, \sigma_\varphi^2), \quad \textit{where } \sigma_\varphi^2 := \sigma^2 \int_D u^2(x)(\mathcal{G}\varphi)^2(x)\, dx. \tag{5-5}$$

Proof. Moving the operator $\mathcal{G}_\varepsilon$ to φ, we have

$$-\frac{1}{\sqrt{\varepsilon^d}}(\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon, \varphi)_{L^2} = -\frac{1}{\sqrt{\varepsilon^d}} \int_D \nu\Big(\frac{x}{\varepsilon}\Big) v^\varepsilon(x) \psi^\varepsilon(x)\, dx,$$

where $\psi^\varepsilon = \mathcal{G}_\varepsilon \varphi$. Let $I_1^\varepsilon[\varphi]$ denote the random variable above. Set $\psi = \mathcal{G}\varphi$ and introduce

$$J_1^\varepsilon[\varphi] := -\frac{1}{\sqrt{\varepsilon^d}} \int_D \nu\Big(\frac{x}{\varepsilon}\Big) u(x) \psi(x)\, dx. \tag{5-6}$$

Since $\nu(x,\omega)$ is a stationary ergodic random field that has short-range correlation, $u\psi \in L^2(D)$, we apply the well-known functional central limit theorem (see, e.g., [Bal 2008, Theorem 3.8]) and obtain

$$J_1^\varepsilon[\varphi] \xrightarrow{\text{distribution}} I_1[\varphi] := \sigma\left(\int_D G(x,y)u(y)\,dW(y),\varphi\right)_{L^2} \sim \mathcal{N}\left(0,\sigma^2\int_D (u(y)\psi(y))^2\,dy\right). \quad (5\text{-}7)$$

The last relation $\sim$ above means equal in law. We note that

$$\mathbb{E}\left|J_1^\varepsilon[\varphi]-I_1^\varepsilon[\varphi]\right|^2 = \frac{1}{\varepsilon^d}\mathbb{E}\left(\int_D \nu\left(\frac{x}{\varepsilon}\right)(v^\varepsilon\psi^\varepsilon-u\psi)\,dx\right)^2 \le C\|v^\varepsilon\psi^\varepsilon-u\psi\|_{L^2}^2,$$

and from periodic homogenization theory, we have $v^\varepsilon \to u$ in L^2, $\psi^\varepsilon\to\psi$ in L^2, as $\varepsilon\to 0$; moreover, v^ε and ψ^ε are bounded in L^∞ since $H^2(D)$ is embedded in $L^\infty(D)$ for $d=2,3$. As a consequence, the right-hand side above converges to zero as $\varepsilon\to 0$. As a result,

$$I_1^\varepsilon[\varphi] = J_1^\varepsilon[\varphi] + (I_1^\varepsilon[\varphi]-J_1^\varepsilon[\varphi])$$

is the sum of a term that converges in distribution to $I_1[\varphi]$ and a term that converges to zero in $L^2(\Omega)$. The desired result follows immediately. □

Finally, we collect the facts obtained above to give a proof of Theorem 2.4(i).

Proof of Theorem 2.4(i). Owing to Lemma 4.2 and Lemma 4.4, for $d=2,3$, we have

$$\mathbb{E}\left\|\varepsilon^{-\frac{d}{2}}\left(\mathcal{G}_\varepsilon\nu_\varepsilon\mathcal{G}_\varepsilon\nu_\varepsilon v^\varepsilon - \mathbb{E}\mathcal{G}_\varepsilon\nu_\varepsilon\mathcal{G}_\varepsilon\nu_\varepsilon v^\varepsilon\right)+\varepsilon^{-\frac{d}{2}}\left(\mathcal{G}_\varepsilon\nu_\varepsilon\mathcal{G}_\varepsilon\nu_\varepsilon(u^\varepsilon-v^\varepsilon)-\mathbb{E}\mathcal{G}_\varepsilon\nu_\varepsilon\mathcal{G}_\varepsilon\nu_\varepsilon(u^\varepsilon-v^\varepsilon)\right)\right\|_{L^2}\le\varepsilon^{\frac{d}{2}}.$$

By Chebyshev's inequality, these two terms, as random elements of $L^2(D)$, converge in probability to the zero function. It follows that the limiting distribution of $\varepsilon^{-d/2}(u^\varepsilon-\mathbb{E}\,u^\varepsilon)$ is given by that of the leading term $X^\varepsilon(\omega) := -\varepsilon^{-d/2}\mathcal{G}_\varepsilon\nu^\varepsilon v^\varepsilon$.

Let X be the right-hand side of (2-15). It is a random element of $L^2(D)$ defined on some probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$ on which the Wiener process $W(y,\tilde\omega)$ is defined. Let the distribution of X be P^X and its characteristic function be ϕ^{P^X}. We note that, for any $\varphi\in L^2(D)$, the inner product (X,φ) has Gaussian distribution $\mathcal{N}(0,\sigma_\varphi^2)$, with σ_φ^2 defined in (5-5). Indeed,

$$\mathbb{E}^{P^X}(X,\varphi) = \sigma\,\mathbb{E}^{P^X}\int_D\left(\int_D G(x,y)\varphi(x)\,dx\right)u(y)\,dW(y) = 0,$$

and

$$\mathbb{E}^{P^X}(X,\varphi)^2 = \sigma^2\,\mathbb{E}^{P^X}\left(\int_D\left(\int_D G(x,y)\varphi(x)\,dx\right)u(y)\,dW(y)\right)^2 = \sigma^2\int_D(\mathcal{G}\varphi)^2u^2\,dy.$$

This shows $(X,\varphi)\sim\mathcal{N}(0,\sigma_\varphi^2)$ in law. By Lemma 5.3, for any fixed $\varphi\in L^2(D)$, the random variable (X^ε,φ) converges in distribution to $\mathcal{N}(0,\sigma_\varphi^2)$. This shows that, as mentioned in Remark 2.2, the characteristic function of the law of X^ε converges to that of X. In view of Lemma 5.1 and Theorem A.1, the distribution of $\{X^\varepsilon\}_{\varepsilon\in(0,1)}$ in $L^2(D)$ is tight as well. Consequently, by applying Theorem 2.1, we complete the proof of Theorem 2.4(i). □

5B. ***Limiting distribution in $H^{-1}(D)$ for dimensions four and five.*** For dimension $d \geq 4$, we do not expect $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}\, u^\varepsilon)$ to converge in distribution in $L^2(D)$, because as shown in (2-13), the fluctuations scale like $\varepsilon^2 |\log \varepsilon|^{1/2}$ for $d = 4$, and scale like ε^2 for $d \geq 5$. In both cases, the scaling is much stronger than $\varepsilon^{d/2}$. Nevertheless, we prove that convergence in law in $H^{-1}(D)$ holds.

As before, the key step is to show that the probability measure in $H^{-1}(D)$ of the scaled leading term $\{X^\varepsilon\} := -\varepsilon^{-d/2}\mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon$ in the expansion (4-11) is tight, and to show that the characteristic function of this measure converges.

Let us first address the characteristic function ϕ^{X^ε}. We note that $L^2(D)$ is naturally embedded to $H^{-1}(D)$. For any $f \in L^2$, the linear form $L_f : H_0^1(D) \to \mathbb{R}$ given by $L_f(\psi) = (f, \psi)$ is clearly an element of $H^{-1}(D)$, and

$$\|L_f\|_{H^{-1}(D)} = \sup_{\psi \in H_0^1(D),\ \|\psi\|_{H^1} \leq 1} L_f(\psi) \leq \|f\|_{L^2}.$$

We henceforth identify $L_f \in H^{-1}(D)$ with f when $f \in L^2(D)$. For any $\ell \in H^{-1}(D)$, let l be the element in $H_0^1(D)$ that is related to ℓ by a Riesz representation. Then we have

$$(f, \ell)_{H^{-1}(D)} = L_f(l) = (f, l).$$

That is, the $H^{-1}(D)$ inner product of $f \in L^2(D)$ with ℓ is the same as the L^2 inner product of f with l. As a result, Remark 2.2 applies for distribution on $H^{-1}(D)$: to show $\phi^{P^{X^\varepsilon}}$ converges to ϕ^{P^X} as characteristic functions of distributions in $H^{-1}(D)$, it suffices to prove $(X^\varepsilon, h) \to (X, h)$ in distribution as random variables for each fixed $h \in L^2(D)$.

Now we address the tightness of the measures of $\{X^\varepsilon\}$. Our strategy is to control the mean of $\|X^\varepsilon\|_{H^{-s}(D)}$ for some $s \in (0, 1)$ and then apply Theorem A.2. To this purpose, we first observe that $X^\varepsilon \in L^2(D)$ and hence $X^\varepsilon \in H^{-s}(D)$ if we set

$$X^\varepsilon : H_0^s(D) \to \mathbb{R} \quad \text{by} \quad X^\varepsilon(h) = \int_D X^\varepsilon(x) h(x)\, dx. \tag{5-8}$$

For any $h \in H_0^s(D)$, the above clearly defines a continuous linear functional. Moreover, if we identify the function X^ε with its extension to $\mathbb{R}^d$ by zero outside D, the above also defines an element in $H^{-s}(\mathbb{R}^d)$. Since ∂D is regular (as a matter of fact, a $C^{0,1}$-boundary is sufficient), any $h \in H_0^s(D)$ can be extended continuously to $Eh \in H^s(\mathbb{R}^d)$, which satisfies $\|Eh\|_{H^s(\mathbb{R}^d)} \leq T\|h\|_{H^s(D)}$; see [Di Nezza et al. 2012, Theorem 5.4]; by duality, $H^{-s}(\mathbb{R}^d)$ is continuously embedded in $H^{-s}(D)$. If fact, we have

$$\begin{aligned}\|X^\varepsilon\|_{H^{-s}(D)} &:= \sup_{w \in H_0^s(D), \|w\|_{H^s(D)} \leq 1} (X^\varepsilon, w)_{L^2} = \sup_{w \in H^s(D), \|w\|_{H_0^s(D)} \leq 1} (X^\varepsilon, Ew)_{L^2} \\ &\leq \sup_{v \in H^s(\mathbb{R}^d), \|v\|_{H^s(\mathbb{R}^d)} \leq T} (X^\varepsilon, v)_{L^2} \leq T\|X^\varepsilon\|_{H^{-s}(\mathbb{R}^d)}.\end{aligned} \tag{5-9}$$

We note that $\|X^\varepsilon\|_{H^{-s}(\mathbb{R}^d)}$ can be calculated using the formula (5-2).

Consider, for each fixed $y \in D$, the Green's function $G_\varepsilon(\cdot, y)$ for the Dirichlet problem (3-7). Extend $G_\varepsilon(\cdot, y)$ to $\mathbb{R}^d$ by zero outside D, and let G_ε^y denote the extended function. Then G_ε^y defines naturally a linear form on $H^s(\mathbb{R}^d)$ by

$$\begin{aligned} G_\varepsilon^y &: H^s(\mathbb{R}^d) \to \mathbb{R}, \\ h &\mapsto G_\varepsilon^y(h) := \int_{\mathbb{R}^d} G_\varepsilon^y(x)h(x)\,dx = \int_D G_\varepsilon(x,y)h(x)\,dx, \end{aligned} \tag{5-10}$$

provided the integral is finite. Since $\mathcal{G}_\varepsilon$ is self-adjoint and by the Green's function representation, $w(y) := G_\varepsilon^y(h)$ is the solution to the Dirichlet problem $\mathcal{L}_\varepsilon w = h$ on D, with zero boundary condition. Note that the restriction of h on D is in $H^s(D)$. Invoking elliptic regularity, we find that w is bounded in $H^{s+2}(D)$. Let $s \in (0,1)$ if $d = 4$ and $s \in \left(\frac{1}{2}, 1\right)$ if $d = 5$; then by the embedding theorem of fractional Sobolev spaces, $H^{s+2}(D) \subset C^{0,\alpha}(D)$ with $\alpha = s + 2 - \frac{d}{2} \in (0,1)$; see [Grisvard 1985, Theorem 1.4.4.1]. As a result, $|G_\varepsilon^y(h)| \le C\|h\|_{H^s}$, where C only depends on the universal constants and the index s. We hence proved the following fact:

Lemma 5.4. *Assume* (A) *holds and* $\bar{q} \ge 0$. *Identify* $G_\varepsilon(\cdot\,, y)$, *for each fixed* $y \in D$, *with the element in* $H^{-s}(\mathbb{R}^d)$ *defined above. Suppose* $s \in (0,1)$ *for* $d = 4$ *and* $s \in \left(\frac{1}{2}, 1\right)$ *for* $d = 5$. *Then there exists* $C > 0$, *depending only on universal parameters and* s, *such that*

$$\|G_\varepsilon(\cdot\,, y)\|_{H^{-s}(\mathbb{R}^d)} \le C. \tag{5-11}$$

Using this fact and the Fourier transform formula for the $H^{-s}(\mathbb{R}^d)$-norm, we can prove the following control of $\|X^\varepsilon\|_{H^{-s}(\mathbb{R}^d)}$ which, together with Theorem A.2, yields the tightness of $\{X^\varepsilon\}$.

Lemma 5.5. *Suppose that the conditions of Theorem 2.3 are satisfied. Assume further that* $d = 4, 5$. *Let* $s \in (0,1)$ *if* $d = 4$ *and* $s \in \left(\frac{1}{2}, 1\right)$ *if* $d = 5$. *Then there exists a constant* $C > 0$, *depending only on the universal parameters and on* s, *such that*

$$\mathbb{E}\|X^\varepsilon\|_{H^{-s}(D)}^2 \le C. \tag{5-12}$$

Proof. We identify X^ε with the element in $H^{-s}(D) \subset H^{-s}(\mathbb{R}^d)$ defined earlier. In view of (5-9), we have

$$\|X^\varepsilon\|_{H^{-s}(D)}^2 \le C\|X^\varepsilon\|_{H^{-s}(\mathbb{R}^d)} = \frac{1}{\varepsilon^d}\int_{\mathbb{R}^d} |\mathcal{F}X^\varepsilon(\xi)|^2 (1+|\xi|^2)^{-s}\,d\xi,$$

where $\mathcal{F}X^\varepsilon$ denotes the Fourier transform of the (extended) function X^ε. Using the integral representation of X^ε, we rewrite the above as

$$\|X^\varepsilon\|_{H^{-s}}^2 = \frac{1}{\varepsilon^d}\int_{\mathbb{R}^d} d\xi\,(1+|\xi|^2)^{-s}\int_{\mathbb{R}^{2d}} dx\,dy \int_{D^2} dz\,dt\, e^{i\xi\cdot(x-y)} G_\varepsilon(x,z)G_\varepsilon(y,t)\nu^\varepsilon(z)v^\varepsilon(z)\nu^\varepsilon(t)v^\varepsilon(t),$$

where the Green's functions are extended by zero to $\mathbb{R}^d$ for their first variables. Take the expectation in this formula; we have

$$\mathbb{E}\|X^\varepsilon\|_{H^{-s}}^2 = \frac{1}{\varepsilon^d}\int_{\mathbb{R}^{2d}}\left(\int_{\mathbb{R}^{3d}} \frac{e^{i\xi\cdot(x-y)}G_\varepsilon(x,z)G_\varepsilon(y,t)}{(1+|\xi|^2)^s}\,dx\,dy\,d\xi\right) R\left(\frac{z-t}{\varepsilon}\right) v^\varepsilon(z)v^\varepsilon(t)\,dt\,dz.$$

We claim that for any z and t in D,

$$\left|\int_{\mathbb{R}^{3d}} \frac{e^{i\xi\cdot(x-y)}G_\varepsilon(x,z)G_\varepsilon(y,t)}{(1+|\xi|^2)^s}\,dx\,dy\,d\xi\right| \le C. \tag{5-13}$$

Indeed, we recognize the quantity inside the absolute value sign to be

$$\int_{\mathbb{R}^d} \frac{\mathscr{F}G_\varepsilon^z(\xi)\,\overline{\mathscr{F}G_\varepsilon^t(\xi)}}{(1+|\xi|^2)^s} d\xi \le \left(\int_{\mathbb{R}^d} |\mathscr{F}G_\varepsilon^z(\xi)|^2(1+|\xi|^2)^{-s}\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} |\mathscr{F}G_\varepsilon^t(\xi)|^2(1+|\xi|^2)^{-s}\right)^{\frac{1}{2}}.$$

The term on the right-hand side is precisely $\|G_\varepsilon^z\|_{H^{-s}(\mathbb{R}^d)}\|G_\varepsilon^t\|_{H^{-s}(\mathbb{R}^d)}$. In view of Lemma 5.4, we can apply (5-11) to get an upper bound for the quantity above and prove (5-11). Then (5-13) follows, which in turn completes the proof. □

Finally, we conclude this section by collecting the facts above and proving Theorem 2.4(ii).

Proof of Theorem 2.4(ii). Step 1: Limiting distribution of the leading term. In view of Theorem A.1 and Lemma 5.5, the probability measures on $H^{-1}(D)$ induced by $\{X^\varepsilon\}$ are tight. To check the limit of the characteristic functions of $\{P^{X^\varepsilon}\}$, it suffices to prove (2-11). This is done in Lemma 5.3. By Theorem 2.1, we conclude that $X^\varepsilon \to X$ in distribution on $H^{-1}(D)$, where X is defined to be the right-hand side of (2-15).

Step 2: Convergence to zero of the higher-order terms. By Lemma 4.2, and $d = 4, 5$, we see that the second term in $u^\varepsilon - \mathbb{E}\,u^\varepsilon$, i.e., $\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon - \mathbb{E}\,\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon v^\varepsilon$, converges in $L^2(\Omega, L^2(D))$ and hence in $L^2(\Omega, H^{-1}(D))$ to the zero function. Similarly, the remainder term $\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon (u^\varepsilon - v^\varepsilon) - \mathbb{E}\mathcal{G}_\varepsilon \nu^\varepsilon \mathcal{G}_\varepsilon \nu^\varepsilon (u^\varepsilon - v^\varepsilon)$ converges to the zero function in $L^1(\Omega, H^{-1}(D))$. These convergence results are stronger than the mode of convergence in distribution in $H^{-1}(D)$. The proof of Theorem 2.4(ii) is thus complete. □

6. The long-range correlated setting

In this section, we consider the setting where $q(x,\omega)$ has long-range correlations. In this setting, the general central limit theorem (Lemma 5.3) does not hold, and we hence restrict to the special case where q is constructed as a function of Gaussian random fields. Limiting theorems in the spirit of Lemma 5.3 are then obtained from Gaussian computations; see, e.g., [Bal et al. 2008; 2012].

Long-range correlated potentials constructed from Gaussian fields. Let $q(x,\omega) = \bar{q} + \nu(x,\omega)$ with $\bar{q}$ a nonnegative constant; we assume:

(L1) $\nu(x,\omega) = \Phi(g(x))$, and $g(x,\omega)$ is a centered stationary Gaussian random field with unit variance. Furthermore, the correlation function of $g(x,\omega)$ has heavy tail. That is, for some positive constant κ_g and some real number $\alpha \in (0,d)$,

$$R_g(x) := \mathbb{E}\{g(y,\omega)g(y+x,\omega)\} \sim \kappa_g |x|^{-\alpha} \quad \text{as } |x| \to \infty. \tag{6-1}$$

(L2) The function $\Phi : \mathbb{R} \to \mathbb{R}$ satisfies $-\bar{q} \le \Phi \le M - \bar{q}$, and has Hermite rank one, i.e.,

$$\int_{\mathbb{R}} \Phi(s)e^{-\frac{s^2}{2}}\,ds = 0, \quad V_1 := \int_{\mathbb{R}} s\Phi(s)e^{-\frac{s^2}{2}}\,ds \ne 0. \tag{6-2}$$

(L3) The Fourier transform $\widehat{\Phi}$ of the function Φ satisfies

$$\int_{\mathbb{R}} |\widehat{\Phi}(\xi)|\big(1+|\xi|^3\big) < \infty. \tag{6-3}$$

We henceforth refer to the above conditions together as (L).

The assumption (L2) makes $\nu(x,\omega) = \Phi(g(x,\omega))$ mean zero, and the bounds on Φ ensure that $0 \le q(x,\omega) \le M$, which is (2-4). From the above construction, we check that $\nu(x,\omega)$ is stationary ergodic and has a long-range correlation function that decays like $\kappa|x|^{-\alpha}$, where $\kappa = V_1^2\kappa_g$; see Lemma A.3 for the details. Assumption (L3) allows one to derive a (nonasymptotic) estimate (see Lemma A.4 in the Appendix) for the fourth-order moments of $\nu(x,\omega)$. Universal constants in the long-range correlation setting may depend on α, R_g, κ_g, Φ and κ.

For the scaling of the homogenization error, we have the following analogue of Theorem 2.3. We focus on the error $u^\varepsilon - \mathbb{E}u^\varepsilon$ because, as seen earlier, the main contribution to the deterministic error $\mathbb{E}u^\varepsilon - u$ comes from the periodic oscillation in the diffusion coefficients, and Theorem 3.3(i) holds independent of the correlation length of $\nu(x,\omega)$.

Theorem 6.1. *Let $D \subset \mathbb{R}^d$ be an open bounded $C^{1,1}$-domain, u^ε and u be the solutions to* (1-1) *and* (1-2) *respectively. Suppose that* (A), (P) *and* (L) *hold and $f \in L^2(D)$. Then, there exists positive constant C, which depends only on the universal parameters, such that if $2 \le d \le 5$ and $0 < \alpha < d$ or $6 \le d \le 7$ and $0 < \alpha < 6$,*

$$\mathbb{E}\|u^\varepsilon - \mathbb{E}u^\varepsilon\|_{L^2} \le \begin{cases} C\varepsilon^{\frac{\alpha}{2}\wedge 2}\|f\|_{L^2} & \text{if } d \ne 4, \\ C\varepsilon^{\frac{\alpha}{2}}\|f\|_{L^2} & \text{if } d = 4. \end{cases} \tag{6-4}$$

Moreover, for any $\varphi \in L^2(D)$, with $2 \le d \le 7$ and $0 < \alpha < d$,

$$\mathbb{E}\big|(u^\varepsilon - \mathbb{E}u^\varepsilon, \varphi)_{L^2}\big| \le C\varepsilon^{\frac{\alpha}{2}}\|\varphi\|_{L^2}\|f\|_{L^2}. \tag{6-5}$$

This result shows that the random fluctuation $u^\varepsilon - \mathbb{E}u^\varepsilon$ caused by the long-range correlated random potential scales like $\varepsilon^{(\alpha\wedge 4)/2}$ in the energy norm, and scales like $\varepsilon^{\alpha/2}$ with respect to the weak topology. Since $\alpha < d$, we note that the random fluctuation in this setting is larger than the case of short-range correlated potential. We mention that if $\alpha < 2$, then the random fluctuations dominate the deterministic fluctuation caused by the periodic diffusion.

The next result exhibits the limiting law of the rescaled random fluctuation $\varepsilon^{-\alpha/2}(u^\varepsilon - \mathbb{E}u^\varepsilon)$. In the presentation, we define formally $W^\alpha(dy)$ as $\dot{W}^\alpha(y)\,dy$; here $\dot{W}^\alpha(y)$ is a centered stationary Gaussian random field with covariance function $\boldsymbol{E}(\dot{W}^\alpha(x)\dot{W}^\alpha(y)) = \kappa|x-y|^{-\alpha}$, where $\boldsymbol{E}$ denotes the expectation with respect to the distribution of $\dot{W}^\alpha$. Here, $\kappa = \kappa_g V_1^2 > 0$, where κ_g and V_1 are defined in (6-1) and (6-2).

Theorem 6.2. *Suppose that the assumptions in Theorem 6.1 hold. Let κ be defined as in* (6-2) *and $G(x,y)$ be the Green's function of* (1-2). *Let $W^\alpha(dy)$ be defined as above. Then*

(i) *For $d = 2, 3$, as $\varepsilon \to 0$,*

$$\frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\sqrt{\varepsilon^\alpha}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sqrt{\kappa}\int_D G(x,y)u(y)W^\alpha(dy) \quad \text{in } L^2(D). \tag{6-6}$$

(2) *For $d = 4, 5$, as $\varepsilon \to 0$, the above holds as convergence in law in $H^{-1}(D)$.*

Remark 6.3. The right-hand side of (6-6) is an integral with respect to the multiparameter Gaussian random processes W^α; we refer to [Khoshnevisan 2002] for the theory. Let X denote the result of the integral. When $d \le 4$, $\alpha < 4$, the Green's function $G(x,\cdot)$ is in $L^{d/(d-\alpha/2)}$ and X is a random element in $L^2(D)$. In general, X is understood through the Fourier transform of its distribution. Given $h^* \in H^{-1}(D)$, the function $\phi^{P^X}(h^*)$ is defined to be $\boldsymbol{E}\exp\bigl(i\sqrt{\kappa}\int_D \langle G(\cdot\,,y), h^*(\cdot)\rangle u(y) W^\alpha(dy)\bigr)$. In particular, for any fixed positive integer N and functions $\{\varphi_i : 1 \le i \le N\}$ in $L^2(D)$, the random variables $I_i := \langle X, \varphi_i\rangle = \sqrt{\kappa}\int_D \langle G(\cdot\,,y), \varphi_i(\cdot)\rangle u(y) W^\alpha(dy)$, with $i = 1,\dots,N$, are joint Gaussian, centered and have covariance matrix $\Sigma_{ij} := \boldsymbol{E}(I_i I_j)$ given by

$$\Sigma_{ij} := \kappa \int_{D^2} \frac{(u\mathcal{G}\varphi_i)(y)(u\mathcal{G}\varphi_j)(z)}{|y-z|^\alpha}\, dy\, dz. \tag{6-7}$$

We will not present the proofs of the results above here, but they can be found in a longer version of this paper [Jing 2015]. The proofs are again based on the expansion formulas (4-11) and (4-17): the leading term has mean zero and contributes to the limiting law; the other terms have larger mean but smaller variance and, after the mean is removed, do not contribute to the limiting law. The main difference in the analysis of the long-range correlation setting is as follows. Firstly, to estimate integrals of $R^\varepsilon(x)$, because $R(x)$ is not integrable, we cannot expect to gain a factor of ε^d by scaling the variable in R^ε. Instead, we gain a factor of ε^α by using the asymptotic of R^ε outside a $(T\varepsilon)$-ball; see Lemma A.3. Secondly, to control fourth-order moments of ν, Lemma 4.3 is no longer useful and we use the estimate in Lemma A.4 instead. In fact, this estimate is less restrictive and, even in the short-range correlation setting, it could be used to replace $\varrho^{1/2}$ in the stronger assumption (S) by ϱ. Last but not least, as mentioned earlier, general central limit theorems, e.g., Lemma 5.3, are not available for the limiting law of the first term in (4-17), and we need to appeal to limit theorems that are special for functions of Gaussian processes.

7. Further discussions

7A. ***An alternative condition for (S).*** In the short-range correlation setting for $\nu(x,\omega)$, we assumed the condition (S). Upon applying Lemma 4.3, we can bound the (partial) fourth-order moment Ψ_ν by the sum of two terms, each consisting of the product of a pair of functions $\vartheta \in L^\infty \cap L^1(\mathbb{R}^d)$. However, as remarked earlier, (S) essentially requires the mixing coefficient $\varrho(r)$, and hence $R(|x|)$, to behave like $o(r^{-2d})$ at infinity, which is much stronger than $R(x)$ being integrable.

We remark that (S) is assumed mainly to simplify the presentation and the $o(r^{-2d})$ decay of ϱ is not necessary. In fact, an alternative assumption used in [Bal and Jing 2011] to control fourth-order moments is: there exists $\vartheta : \mathbb{R}^d \to \mathbb{R}_+$ in $L^1 \cap L^\infty(\mathbb{R}^d)$ such that (A-6) holds. This is clearly a much more general assumption, and it is satisfied if $\nu(x,\omega) = \Phi(g(x,\omega))$, with Φ satisfying (L2) and (L3), and $g(x,\omega)$ a centered stationary Gaussian random field with correlation function $R_g = o(|x|^{-d})$ as $|x| \to \infty$.

The conclusions of Theorem 2.4 still hold if (S) is replaced by the above alternative assumption. Indeed, we only need to modify the control of Ψ_ν in the proof of Lemma 4.2 and in Section 4B. For instance, in the first inequality in the proof of Lemma 4.2, we have more but finitely many integrals instead of two on the right-hand side. Nevertheless, in all of these integrals, at most one of the functions ϑ has the same variable as one of the Green's function, and all of them can be controlled. We refer to [Jing 2015, Section 6] for the details.

7B. ***Comparison with the case of nonoscillatory diffusion.*** The main results of this paper show that the framework developed in [Bal 2008; Bal and Jing 2011; Bal et al. 2012], in the setting of a nonoscillatory differential operator with oscillatory random potential, applies even when the differential operator is also oscillatory, as long as we have uniform-in-ε control of the Green's functions and their gradients, i.e., (3-11) and (3-12), and provided that there is no random correlation between the diffusion coefficients and the potential. At the formal level, there is no difference in the proof, and the usual strategy using (truncated) series expansion applies. However, the role played by the oscillatory diffusion coefficients becomes prominent in getting the tightness of the measures of $\{X^\varepsilon\}$.

Let us recall the previous method used for tightness in the setting of a nonoscillatory differential operator. Set

$$\mathcal{L} := -\sum_{i,j=1}^{d} \bar{a}_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \bar{q},$$

and consider the Dirichlet problem $(\mathcal{L} + \nu^\varepsilon)u^\varepsilon = f$ in D with zero boundary condition. Then u^ε homogenizes to u, the solution of (1-2). As in [Bal et al. 2012], the limiting distribution of $\varepsilon^{-d/2}(u^\varepsilon - \mathbb{E}u^\varepsilon)$, say, in the short-range correlation setting, is characterized by that of $X^\varepsilon = -\varepsilon^{-d/2}\mathcal{G}\nu^\varepsilon u$. To prove tightness of the measures $\{X^\varepsilon\}$ in $L^2(D)$, the strategy of [Bal et al. 2012] is to use the spectral representation of $L^2(D)$. Note that $\mathcal{L}$ is formally self-adjoint and its inverse, i.e., $\mathcal{G}$, is compact on $L^2(D)$. Hence, $\mathcal{L}$ admits real eigenvalues $\{\lambda_k\}_{k=1}^\infty$ such that,

$$0 \le \bar{q} < \lambda_1 \le \lambda_2 \le \cdots, \quad \lambda_k \to \infty \quad \text{as} \quad k \to \infty,$$

and eigenfunctions $\{\phi_k\}_{k=1}^\infty$, with $\|\phi_k\|_{L^2} = 1$, such that

$$\begin{cases} \mathcal{L}\phi_k = \lambda_k \phi_k & \text{in } D, \\ \phi_k = 0 & \text{on } \partial D. \end{cases}$$

Moreover, $\{\phi_k\}$ form an orthonormal basis of $L^2(D)$ and we have the following representation of the space $\mathcal{H}^0(D) = L^2(D)$ and the Sobolev space $\mathcal{H}^1(D) = H_0^1(D)$; see [Evans 1998, Section 6.5]: for $s = 0, 1$,

$$\mathcal{H}^s(D) = \overline{\left\{ f \in C^\infty(D) : \sum_{k=1}^\infty (f, \phi_k)_{L^2}^2 \lambda_k^s < \infty \right\}} \quad \text{and} \quad \|v\|_{\mathcal{H}^s}^2 := \sum_{k=1}^\infty (f, \phi_k)_{L^2}^2 \lambda_k^s. \tag{7-1}$$

A natural criterion for tightness of (the measures of) $\{X^\varepsilon\}$ is that their measures do not scatter to higher and higher modes. More precisely, let P_N denote the projection operator in $L^2(D)$ to the space $W_N := \operatorname{span}\{\phi_1, \dots, \phi_N\}$ spanned by the first N modes. Then $\{X^\varepsilon\}$ is tight if $\mathbb{E}\|X^\varepsilon\|_{L^2} \le C$ and

$$\lim_{N\to\infty} \sup_{\varepsilon\in(0,1)} \mathbb{E}\|X^\varepsilon - P_N X^\varepsilon\|_{L^2} = 0. \tag{7-2}$$

Using the representation formula in (7-1), and the fact that $\mathcal{G}\phi_k = (\lambda_k)^{-1}\phi_k$, we have

$$\mathbb{E}\|X^\varepsilon - P_N X^\varepsilon\|_{L^2}^2 = \frac{1}{\varepsilon^d} \sum_{k=N+1}^\infty \mathbb{E}(\mathcal{G}\nu^\varepsilon u, \phi_k)^2 = \frac{1}{\varepsilon^d} \sum_{k=N+1}^\infty \frac{1}{\lambda_k^2} \mathbb{E}(\nu^\varepsilon u, \phi_k)^2.$$

As in Section 4B, we have $\sup_{\varepsilon\in(0,1)} \sup_k \mathbb{E}(\nu^\varepsilon u, \phi^k)^2 \le C$. In view of Weyl's asymptotic formula for the eigenvalues, $\lambda_k \approx k^{2/d}$ for k large, we conclude that

$$\sup_{\varepsilon\in(0,1)} \mathbb{E}\|X^\varepsilon - P_N X^\varepsilon\|_{L^2}^2 \lesssim \sum_{k=N+1}^{\infty} \frac{1}{\lambda_k^2} \lesssim \sum_{k=N+1}^{\infty} \frac{1}{k^{\frac{4}{d}}}.$$

Hence, for $d = 2, 3$, we obtain tightness of $\{X^\varepsilon\}$ for free, as byproduct of the analysis in Section 4B.

In the setting of this paper, $\mathcal{L}$ above is replaced by $\mathcal{L}_\varepsilon$, defined in (3-1). The above approach for tightness fails completely. On the one hand, if we replace the eigenpairs $(\lambda_k, \phi_k)_k$ by $(\lambda_k^\varepsilon, \phi_k^\varepsilon)_k$, where the latter solve the eigenvalue problems associated to $\mathcal{L}_\varepsilon$, then instead of (7-2), we obtain

$$\lim_{N\to\infty} \sup_{\varepsilon\in(0,1)} \mathbb{E}\|X^\varepsilon - P_N^\varepsilon X^\varepsilon\|_{L^2} = 0,$$

where P_N^ε is the projection to $W_N^\varepsilon := \operatorname{span}\{\phi_1^\varepsilon, \dots, \phi_N^\varepsilon\}$. This is useless because, a priori, the basis $(\phi_k^\varepsilon)_k$ changes with ε, and it is not clear that the union (over $\varepsilon \in (0,1)$) of unit balls in W_N^ε is still compact for all N. On the other hand, if we fix a spectral representation, say, using $(\lambda_k, \phi_k)_k$ defined before, then we no longer have the relation $\mathcal{G}_\varepsilon \phi_k = (\lambda_k)^{-1}\phi_k$. It is not difficult to check that $\|\nabla \mathcal{G}_\varepsilon \phi_k\|_{L^2} \approx 1/\sqrt{\lambda_k}$ and this estimate is sharp. An application of the Poincaré inequality yields that $\|\mathcal{G}_\varepsilon \phi_k\|_{L^2} \le C/\sqrt{\lambda_k} \approx k^{-1/d}$, with C uniform in ε and k. It is not clear at all how to improve this estimate. Consequently, in view of the estimate on I_1^ε in Section 4B, we have

$$\sup_{\varepsilon\in(0,1)} \mathbb{E}\|X^\varepsilon - P_N X^\varepsilon\|_{L^2}^2 = \frac{1}{\varepsilon^d} \sum_{k=N+1}^{\infty} \mathbb{E}(\nu^\varepsilon u, \mathcal{G}_\varepsilon \phi_k)^2 \le \sum_{k=N+1}^{\infty} C\|\mathcal{G}_\varepsilon \phi_k\|_{L^2}^2 \approx \sum_{k=N+1}^{\infty} \frac{1}{\lambda_k} \approx \sum_{k=N+1}^{\infty} \frac{1}{k^{\frac{2}{d}}}.$$

This fails to show (7-2) or the tightness of $\{X^\varepsilon\}$, even for $d = 2$.

In view of the analysis above, we find that the above approach for tightness, which is natural for nonoscillatory differential operators, fails completely in the presence of fast oscillations in the diffusion coefficients. The new approach used in Section 5 is necessary and more stable.

Appendix: Some technical results

Tightness criteria for probability measures in functional spaces. We first present a tightness criterion for the probability measures $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ on $L^2(D)$ induced by $\{X^\varepsilon(\cdot, \omega)\}$ that are random elements in $H_0^s(D) \subset L^2(D)$, with $s \in (0, 1]$.

Theorem A.1 (tightness in $L^2(D)$). *Let $\{X^\varepsilon(\cdot,\omega)\}_{\varepsilon\in(0,1)}$ be a family of random fields on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $X^\varepsilon(\cdot,\omega) \in H_0^s(D)$ for some $0 < s \le 1$, for each fixed $\varepsilon \in (0,1)$ and $\omega \in \Omega$. Suppose there exists $C > 0$, independent of ε and ω, such that*

$$\mathbb{E}\|X^\varepsilon\|_{H^s} \le C. \tag{A-1}$$

Then the family of probability measures $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ on $L^2(D)$ is tight.

Proof. By assumption, P^{X^ε} concentrates on the subspace $H_0^s(D)$. For any fixed $\delta > 0$, set $M_\delta = C\delta^{-1}$ and define

$$\mathcal{A}_\delta = \{ f \in H_0^s(D) : \|f\|_{H^s} \le M_\delta \}.$$

Clearly, $\mathcal{A}_\delta$ is closed and bounded in $H_0^s(D)$. In light of the fact that the embedding $H_0^s(D) \hookrightarrow L^2(D)$ is compact [Palatucci et al. 2013], we note that $\mathcal{A}_\delta$ is a compact set of $L^2(D)$. Now for any fixed $\varepsilon \in (0,1)$, applying Chebyshev's inequality, we find

$$\begin{aligned} P^{X^\varepsilon}(\mathcal{A}_\delta) = \mathbb{P}\big(\{X^\varepsilon \in H_0^s(D),\ \|X^\varepsilon\|_{H^s} \le M_\delta\}\big) &= 1 - \mathbb{P}(\{\|X^\varepsilon\|_{H^s} > M_\delta\}) \\ \ge 1 - \frac{\mathbb{E}\|X^\varepsilon\|_{H^s}}{M_\delta} \ge 1 - \frac{C}{M_\delta} = 1 - \delta. \end{aligned}$$

Since δ and ε are arbitrary, the above shows that $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ is tight. ☐

Next we give a similar tightness criterion for probability measures $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ on $H^{-1}(D)$ induced by $\{X^\varepsilon(\cdot,\omega)\}$ which belong to a smoother space.

Theorem A.2 (tightness in $H^{-1}(D)$). *Let $\{X^\varepsilon(\cdot,\omega)\}_{\varepsilon\in(0,1)}$ be a family of random fields on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $X^\varepsilon(\cdot,\omega) \in H^{-s}(D)$ for some $0 \le s < 1$, for each fixed $\varepsilon \in (0,1)$ and $\omega \in \Omega$. Suppose there exists a constant $C > 0$, independent of ε and ω, such that*

$$\mathbb{E}\|X^\varepsilon\|_{H^{-s}} \le C. \tag{A-2}$$

Then the probability measures $\{P^{X^\varepsilon}\}_{\varepsilon\in(0,1)}$ on $H^{-1}(D)$ are tight.

Proof. Since D is a bounded open set with regular boundary, the embedding $H_0^1(D) \hookrightarrow H_0^s(D)$, for any $0 \le s < 1$, is compact [Grisvard 1985, Theorem 1.4.3.2]. By duality, the embedding $H^{-s}(D) \hookrightarrow H^{-1}(D)$ is also compact. The rest of the proof is exactly the same as in the proof of the theorem above. ☐

Functions of long-range correlated Gaussian random fields. Here we record some results for the random potential $\nu(x,\omega) = \Phi(g(x,\omega))$ that is constructed in (L). In particular, we express the asymptotic behavior of its correlation function $R(x)$, and derive a (partial) fourth-order moment for ν.

Autocorrelation function of the long-range model.

Lemma A.3. *Assume* (L1) *and* (L2) *hold and let $\nu(x,\omega)$ be as constructed there. Set $V_1 = \mathbb{E}\{g_0\Phi(g_0)\}$, where g_x is the underlying Gaussian random field in* (L). *Then there exist constants $T, C > 0$, depending only on the universal parameters, such that the autocorrelation function $R(x)$ of q satisfies*

$$|R(x) - V_1^2 R_g(x)| \le C R_g^2(x) \quad \textit{for all } |x| \ge T, \tag{A-3}$$

where R_g is the correlation function of g. Further,

$$\big|\mathbb{E}\{g(y)q(y+x)\} - V_1 R_g(x)\big| \le C R_g^2(x) \quad \textit{for all } |x| \ge T. \tag{A-4}$$

The proof of this result can be found in [Bal et al. 2008; 2012]. It says that $\nu(x,\omega)$ inherits the heavy tail from the underlying Gaussian random field. The next result describes estimates on the integrals of R, possibly against some potential function that has singularity at the origin.

Fourth-order moments of $\nu(x,\omega)$. Finally, we present a nonasymptotic estimate for the four-moments of $\nu(x,\omega)$ constructed in (L1) and (L2), with the additional assumption (L3). In the following, we denote by $\mathcal{U}$ the collections of two pairs of unordered numbers in the set $\{1,2,3,4\}$,

$$\mathcal{U} := \big\{p = \{(p(1), p(2)), (p(3), p(4))\} : p(i) \in \{1,2,3,4\},\ p(1) \neq p(2),\ p(3) \neq p(4)\big\}. \qquad \text{(A-5)}$$

As members in a set, the pairs $(p(1), p(2))$ and $(p(3), p(4))$ are required to be distinct; however, the two pairs can have one common index. There are three elements in $\mathcal{U}$ that collect all four numbers. They are precisely $\{(1,2),(3,4)\}$, $\{(1,3),(2,4)\}$ and $\{(1,4),(2,3)\}$. Let $\mathcal{U}_*$ denote the subset formed by these three elements, and let $\mathcal{U}^*$ be its complement.

Lemma A.4. *Assume* (L) *holds and let* $\nu(x,\omega)$ *be as constructed there. Then there exists* $\vartheta : \mathbb{R}^d \to \mathbb{R}_+$, *bounded and satisfying* $\vartheta(x) \sim |x|^{-\alpha}$ *as* $|x| \to \infty$, *and some* $C > 0$, *depending only on the universal parameters*, *such that for any four points* $\{x_i \in \mathbb{R}^d : 1 \le i \le 4\}$,

$$\left| \mathbb{E} \prod_{i=1}^{4} \nu(x_i) - \sum_{p \in \mathcal{U}_*} R(x_{p(1)} - x_{p(2)}) R(x_{p(3)} - x_{p(4)}) \right| \le C \sum_{p \in \mathcal{U}^*} \vartheta(x_{p(1)} - x_{p(2)}) \vartheta(x_{p(3)} - x_{p(4)}). \quad \text{(A-6)}$$

We refer to [Bal and Jing 2011, Proposition 4.1] for the proof of this result. In particular, ϑ above can be chosen as the autocorrelation function $R(x)$ of $\nu(x,\omega)$. As discussed earlier, (A-6) can be viewed as an alternative for the estimates in Lemma 4.3.

Acknowledgments

The author would like to thank the referee for the valuable comments that helped to improve the presentation of the paper. This work is supported in part by NSF grant DMS-1515150.

References

[Armstrong and Shen 2015] S. N. Armstrong and Z. Shen, "Lipschitz estimates in almost-periodic homogenization", *Comm. Pure Appl. Math* (online publication October 2015).

[Armstrong and Smart 2014] S. N. Armstrong and C. K. Smart, "Quantitative stochastic homogenization of convex integral functionals", 2014. To appear in *Ann. Sci. Éc. Norm. Supér.* arXiv 1406.0996

[Armstrong et al. 2015] S. N. Armstrong, T. Kuusi, and J.-C. Mourrat, "Mesoscopic higher regularity and subadditivity in elliptic homogenization", preprint, 2015. arXiv 1507.06935

[Avellaneda and Lin 1987] M. Avellaneda and F.-H. Lin, "Compactness methods in the theory of homogenization", *Comm. Pure Appl. Math.* **40**:6 (1987), 803–847. MR 88i:35019 Zbl 0632.35018

[Avellaneda and Lin 1991] M. Avellaneda and F.-H. Lin, "L^p bounds on singular integrals in homogenization", *Comm. Pure Appl. Math.* **44**:8-9 (1991), 897–910. MR 92j:42015 Zbl 0761.42008

[Bal 2008] G. Bal, "Central limits and homogenization in random media", *Multiscale Model. Simul.* **7**:2 (2008), 677–702. MR 2010b:35508 Zbl 1181.35353

[Bal and Jing 2010] G. Bal and W. Jing, "Homogenization and corrector theory for linear transport in random media", *Discrete Contin. Dyn. Syst.* **28**:4 (2010), 1311–1343. MR 2011g:35025 Zbl 1206.35027

[Bal and Jing 2011] G. Bal and W. Jing, "Corrector theory for elliptic equations in random media with singular Green's function: application to random boundaries", *Commun. Math. Sci.* **9**:2 (2011), 383–411. MR 2012m:35009 Zbl 1237.35165

[Bal et al. 2008] G. Bal, J. Garnier, S. Motsch, and V. Perrier, "Random integrals and correctors in homogenization", *Asymptot. Anal.* **59**:1-2 (2008), 1–26. MR 2009m:35019 Zbl 1157.34048

[Bal et al. 2012] G. Bal, J. Garnier, Y. Gu, and W. Jing, "Corrector theory for elliptic equations with long-range correlated random potential", *Asymptot. Anal.* **77**:3-4 (2012), 123–145. MR 2977330 Zbl 1259.35235

[Billingsley 1999] P. Billingsley, *Convergence of probability measures*, 2nd ed., John Wiley & Sons, New York, 1999. MR 2000e:60008 Zbl 0944.60003

[Bourgeat and Piatnitski 1999] A. Bourgeat and A. Piatnitski, "Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator", *Asymptot. Anal.* **21**:3-4 (1999), 303–315. MR 2000i:34104 Zbl 0960.60057

[Di Nezza et al. 2012] E. Di Nezza, G. Palatucci, and E. Valdinoci, "Hitchhiker's guide to the fractional Sobolev spaces", *Bull. Sci. Math.* **136**:5 (2012), 521–573. MR 2944369 Zbl 1252.46023

[Evans 1998] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics **19**, American Mathematical Society, Providence, RI, 1998. MR 99e:35001 Zbl 0902.35002

[Figari et al. 1982] R. Figari, E. Orlandi, and G. Papanicolaou, "Mean field and Gaussian approximation for partial differential equations with random coefficients", *SIAM J. Appl. Math.* **42**:5 (1982), 1069–1077. MR 84j:60070 Zbl 0498.60066

[Gloria and Otto 2014] A. Gloria and F. Otto, "Quantitative results on the corrector equation in stochastic homogenization", 2014. To appear in *J. Eur. Math. Soc.* arXiv 1409.0801

[Griso 2006] G. Griso, "Interior error estimate for periodic homogenization", *Anal. Appl.* (*Singap.*) **4**:1 (2006), 61–79. MR 2007d:35014 Zbl 1098.35016

[Grisvard 1985] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics **24**, Pitman (Advanced Publishing Program), Boston, 1985. MR 86m:35044 Zbl 0695.35060

[Gu and Mourrat 2015] Y. Gu and J. C. Mourrat, "Scaling limit of flucutations in stochastic homogenization", 2015, Available at http://perso.ens-lyon.fr/jean-christophe.mourrat/scale-limit.pdf. To appear in *Multiscale Model. Simul.*

[Hairer et al. 2013] M. Hairer, E. Pardoux, and A. Piatnitski, "Random homogenisation of a highly oscillatory singular potential", *Stoch. Partial Differ. Equ. Anal. Comput.* **1**:4 (2013), 571–605. MR 3327517 Zbl 1323.35230

[Jikov et al. 1994] V. V. Jikov, S. M. Kozlov, and O. A. Oleĭnik, *Homogenization of differential operators and integral functionals*, Springer, Berlin, 1994. MR 96h:35003b

[Jing 2015] W. Jing, "Limiting distribution of elliptic homogenization error with periodic diffusion and random potential (long version)", preprint, 2015. arXiv 1505.02721

[Kenig et al. 2012] C. E. Kenig, F. Lin, and Z. Shen, "Convergence rates in L^2 for elliptic homogenization problems", *Arch. Ration. Mech. Anal.* **203**:3 (2012), 1009–1036. MR 2928140 Zbl 1258.35086

[Kenig et al. 2014] C. E. Kenig, F. Lin, and Z. Shen, "Periodic homogenization of Green and Neumann functions", *Comm. Pure Appl. Math.* **67**:8 (2014), 1219–1262. MR 3225629 Zbl 1300.35030

[Khoshnevisan 2002] D. Khoshnevisan, *Multiparameter processes: an introduction to random fields*, Springer, New York, 2002. MR 2004a:60003 Zbl 1005.60005

[Lieb and Loss 2001] E. H. Lieb and M. Loss, *Analysis*, 2nd ed., Graduate Studies in Mathematics **14**, American Mathematical Society, Providence, RI, 2001. MR 2001i:00001 Zbl 0966.26002

[Marahrens and Otto 2015] D. Marahrens and F. Otto, "On annealed elliptic Green's function estimates", *Math. Bohem.* **140**:4 (2015), 489–506.

[Moskow and Vogelius 1997] S. Moskow and M. Vogelius, "First-order corrections to the homogenised eigenvalues of a periodic composite medium: a convergence proof", *Proc. Roy. Soc. Edinburgh Sect. A* **127**:6 (1997), 1263–1299. MR 99g:35018 Zbl 0888.35011

[Mourrat 2015] J. C. Mourrat, "A tightness criterion in local Hölder spaces of negative regularity", preprint, 2015. arXiv 1502.07335

[Mourrat and Nolen 2015] J. Mourrat and J. Nolen, "Scaling limit of the corrector in stochastic homogenization", submitted, 2015. arXiv 1502.07440

[Mourrat and Otto 2014] J. C. Mourrat and F. Otto, “Correlation structure of the corrector in stochastic homogenization”, 2014, Available at http://perso.ens-lyon.fr/jean-christophe.mourrat/correl.pdf. To appear in *Ann. Probab.*

[Palatucci et al. 2013] G. Palatucci, O. Savin, and E. Valdinoci, “Local and global minimizers for a variational energy involving a fractional norm”, *Ann. Mat. Pura Appl.* (4) **192**:4 (2013), 673–718. MR 3081641 Zbl 1278.82022

[Parthasarathy 1967] K. R. Parthasarathy, *Probability measures on metric spaces*, Prob. Math. Stat. **3**, Academic Press, New York, 1967. MR 37 #2271 Zbl 0153.19101

Received 8 Jun 2015. Revised 8 Oct 2015. Accepted 28 Oct 2015.

WENJIA JING: wjing@math.uchicago.edu
Department of Mathematics, University of Chicago, 5734 South University Avenue, Chicago, IL 60637, United States

ANALYSIS AND PDE
Vol. 9, No. 1, 2016
dx.doi.org/10.2140/apde.2016.9.229

BLOW-UP RESULTS FOR A STRONGLY PERTURBED SEMILINEAR HEAT EQUATION: THEORETICAL ANALYSIS AND NUMERICAL METHOD

VAN TIEN NGUYEN AND HATEM ZAAG

We consider a blow-up solution for a strongly perturbed semilinear heat equation with Sobolev subcritical power nonlinearity. Working in the framework of similarity variables, we find a Lyapunov functional for the problem. Using this Lyapunov functional, we derive the blow-up rate and the blow-up limit of the solution. We also classify all asymptotic behaviors of the solution at the singularity and give precise blow-up profiles corresponding to these behaviors. Finally, we attain the blow-up profile numerically, thanks to a new mesh-refinement algorithm inspired by the rescaling method of Berger and Kohn. Note that our method is applicable to more general equations, in particular those with no scaling invariance.

1. Introduction

We are concerned in this paper with blow-up phenomena arising in the nonlinear heat problem

$$\begin{cases} \partial_t u = \Delta u + |u|^{p-1}u + h(u), \\ u(\cdot\,,0) = u_0 \in L^\infty(\mathbb{R}^n), \end{cases} \tag{1-1}$$

where $u(t) : x \mapsto u(x,t) \in \mathbb{R}$ for $x \in \mathbb{R}^n$ and Δ stands for the Laplacian in $\mathbb{R}^n$. The exponent $p > 1$ is subcritical (which means that $p < (n+2)/(n-2)$ if $n \ge 3$) and h is given by

$$h(z) = \mu \frac{|z|^{p-1}z}{\log^a(2+z^2)} \quad \text{with} \quad a > 0,\ \mu \in \mathbb{R}. \tag{1-2}$$

By standard results, the problem (1-1) has a unique classical solution $u(x,t)$ in $L^\infty(\mathbb{R}^n)$, which exists at least for small times. The solution $u(x,t)$ may develop singularities in some finite time. We say that $u(x,t)$ blows up in a finite time T if $u(x,t)$ satisfies (1-1) in $\mathbb{R}^n \times [0,T)$ and

$$\lim_{t\to T} \|u(t)\|_{L^\infty(\mathbb{R}^n)} = +\infty.$$

T is called the blow-up time of $u(x,t)$. In such a blow-up case, a point $b \in \mathbb{R}^n$ is called a blow-up point of $u(x,t)$ if and only if there exist $(x_n,t_n) \to (b,T)$ such that $|u(x_n,t_n)| \to +\infty$ as $n \to +\infty$.

In the case $\mu = 0$, the equation (1-1) is the semilinear heat equation

$$\partial_t u = \Delta u + |u|^{p-1}u. \tag{1-3}$$

Zaag is supported by the ERC Advanced Grant no. 291214, BLOWDISOL and by the ANR project ANAÉ ref. ANR-13-BS01-0010-03.
MSC2010: primary 35K10; secondary 35K58.
Keywords: blow-up, Lyapunov functional, asymptotic behavior, blow-up profile, semilinear heat equation, lower-order term.

Problem (1-3) has been addressed in different ways in the literature. The existence of blow-up solutions has been proved by several authors (see [Fujita 1966; Levine 1973; Ball 1977]). Consider a solution $u(x,t)$ of (1-3) which blows up at a time T. The very first question to be answered is the blow-up rate, i.e., there are positive constants C_1 and C_2 such that

$$C_1(T-t)^{-\frac{1}{p-1}} \le \|u(t)\|_{L^\infty(\mathbb{R}^n)} \le C_2(T-t)^{-\frac{1}{p-1}} \quad \text{for all } t \in (0,T). \tag{1-4}$$

The lower bound in (1-4) follows by a simple argument based on Duhamel's formula (see [Weissler 1981]). For the upper bound, Giga and Kohn [1987] proved (1-4) for $1 < p < (3n+8)/(3n-4)$ or for nonnegative initial data with subcritical p.

Later, the estimate (1-4) was extended to all subcritical p without assuming nonnegativity for initial data u_0 by Giga, Matsui and Sasayama [Giga et al. 2004a]. The estimate (1-4) is a fundamental step to obtain more information about the asymptotic blow-up behavior, locally near a given blow-up point $\hat{b}$. Giga and Kohn [1989] showed that, for a given blow-up point $\hat{b} \in \mathbb{R}^n$,

$$\lim_{t\to T} (T-t)^{\frac{1}{p-1}} u(\hat{b} + y\sqrt{T-t}, t) = \pm\kappa,$$

where $\kappa = (p-1)^{-1/(p-1)}$, uniformly on compact sets of $\mathbb{R}^n$.

This result was specified by Filippas and Liu [1993] (see also [Filippas and Kohn 1992]) and Velázquez [1992; 1993] (see also [Herrero and Velázquez 1992a; 1992c; 1993]). Using the renormalization theory, Bricmont and Kupiainen [1994] showed the existence of a solution of (1-3) such that

$$\|(T-t)^{\frac{1}{p-1}} u(\hat{b} + z\sqrt{(T-t)|\log(T-t)|}, t) - f_0(z)\|_{L^\infty(\mathbb{R}^n)} \to 0 \quad \text{as } t \to T, \tag{1-5}$$

where

$$f_0(z) = \kappa\left(1 + \frac{p-1}{4p}|z|^2\right)^{-\frac{1}{p-1}}. \tag{1-6}$$

Merle and Zaag [1997] obtained the same result through a reduction to a finite-dimensional problem. Moreover, they showed that the profile (1-6) is stable under perturbations of initial data (see also [Fermanian Kammerer et al. 2000; Fermanian Kammerer and Zaag 2000; Masmoudi and Zaag 2008] for related results).

In the program developed by those authors in the case $\mu = 0$, the invariance of (1-1) under the scaling transformation

$$\lambda \mapsto u_\lambda(\xi,\tau) = \lambda^{\frac{2}{p-1}} u(\lambda\xi, \lambda^2\tau)$$

played a crucial role. Indeed, this property is responsible for having an autonomous equation in similarity variables defined in (1-10) below (see (1-11) below when $\mu = 0$), which helps a lot.

A similar situation is available for the equation

$$\partial_t u = \Delta u + e^u$$

(see [Herrero and Velázquez 1993; Bebernes and Bricher 1992; Bressan 1990; 1992]).

With more general nonlinearities, namely with

$$\partial_t u = \Delta u + f(u) \tag{1-7}$$

with $f(u) \not\equiv |u|^{p-1}u$ and $f(u) \not\equiv e^u$, no result is available on the blow-up behavior. The first example available in the literature goes back to Giga and Kohn [1987], who considered (1-1) with a "weak" perturbation, namely

$$|h(z)| \le M(|z|^q + 1), \quad q \in [1, p). \tag{1-8}$$

They could extend various results from the case $h \equiv 0$.

In our paper, we aim at doing better, by considering "strong" (in comparison with (1-8)) perturbations, namely the case mentioned in (1-2). The resulting nonlinearity is so close to the power law $|u|^{p-1}u$ that it is not a priori clear if the perturbation is able to modify the blow-up behavior of the solution. A subtle point is the following:

When $\mu = 0$, the similarity variables' version of the PDE is autonomous (see (1-11) below with $\mu = 0$), and classical energy methods à la [Levine 1973] give a Lyapunov functional (see (1-16) below) whose role was crucial in the blow-up analysis performed by Giga and Kohn [1987; 1989] and later authors.

When $\mu \neq 0$, it is still possible to use the similarity variables, however, the resulting equation is not autonomous (see (1-11) below). Moreover, the size of the perturbations introduced by the h term is larger than in the "weak" case (1-8) and, more importantly, it is a priori larger than the correction computed for the solution when $\mu = \mathbb{O}(1/s^a)$ with $0 < a < 1$ as shown in Lemma 2.1, versus $1/s$ in the generic case when $\mu = 0$. New ideas are crucially needed, in particular to find a perturbed Lyapunov functional (see Theorem 1.1 below), and to go beyond the too-large perturbation term $1/s^a$ (we linearize around ϕ defined in (1-21)–(1-22) instead of κ).

Because of those difficulties and thanks to our new ideas, we believe that our paper gives a new framework to the study of blow-up for semilinear heat equations of the type (1-7) when the nonlinearity $f(u)$ could lack any scale invariance (exact, or approximate as in this case) at all.

In the case when the function h satisfies

$$|h(z)| \le M\left(\frac{|z|^p}{\log^a(2+z^2)} + 1\right) \quad \text{with} \quad a > 1 \tag{1-9}$$

and $M > 0$, the first author derived the existence of a Lyapunov functional in the similarity variables (1-10) for the problem (1-1), which is a crucial step in deriving the estimate (1-4). He also classified all possible blow-up behaviors of the solution when it approaches to singularity. Here, we aim at extending the results of [Nguyen 2015] to the case $a \in (0, 1]$. As we mentioned above, the first step is to derive the blow-up rate of the blow-up solution. As in [Giga et al. 2004a; Nguyen 2015], the key step is to find a Lyapunov functional in *similarity variables* for (1-1). More precisely, we introduce for all $b \in \mathbb{R}^n$ (b may be a blow-up point of u or not) the following *similarity variables*:

$$y = \frac{x-b}{\sqrt{T-t}}, \quad s = -\log(T-t), \quad w_{b,T} = (T-t)^{\frac{1}{p-1}} u(x,t). \tag{1-10}$$

Hence $w_{b,T}$ satisfies, for all $s \geq -\log T$ and all $y \in \mathbb{R}^n$,

$$\partial_s w_{b,T} = \frac{1}{\rho}\operatorname{div}(\rho\nabla w_{b,T}) - \frac{w_{b,T}}{p-1} + |w_{b,T}|^{p-1}w_{b,T} + e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}w_{b,T}), \tag{1-11}$$

where

$$\rho(y) = \left(\frac{1}{4\pi}\right)^{\frac{n}{2}} e^{-\frac{|y|^2}{4}} \tag{1-12}$$

and

$$|e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}z)| \leq \frac{C_0}{s^a}(|z|^p+1) \quad \text{for all } z \in \mathbb{R} \tag{1-13}$$

for some $C_0 > 0$.

Following the method introduced by Hamza and Zaag [2012a; 2012b] for perturbations of the semilinear wave equation, we introduce

$$\mathcal{J}_a[w](s) = \mathscr{E}[w](s)e^{\frac{\gamma}{a}s^{-a}} + \theta s^{-a}, \tag{1-14}$$

where γ and θ are positive constants, depending only on p, a, μ and n, which will be determined later, and

$$\mathscr{E}[w] = \mathscr{E}_0[w] + \mathcal{I}[w], \tag{1-15}$$

where

$$\mathscr{E}_0[w](s) = \int_{\mathbb{R}^n}\left(\frac{1}{2}|\nabla w|^2 + \frac{1}{2(p-1)}|w|^2 - \frac{1}{p+1}|w|^{p+1}\right)\rho\, dy \tag{1-16}$$

and

$$\mathcal{I}[w](s) = -e^{-\frac{(p+1)}{p-1}s}\int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}}w)\rho\, dy, \quad H(z) = \int_0^z h(\xi)\, d\xi. \tag{1-17}$$

The main novelty of this paper is to allow values of a in $(0, 1]$, which is possible at the expense of taking the particular form (1-2) for the perturbation h. We aim at the following:

Theorem 1.1 (existence of a Lyapunov functional for (1-11)). *Let a, p, n and μ be fixed; consider w a solution of* (1-11). *Then there exist $\hat{s}_0 = \hat{s}_0(a, p, n, \mu) \geq s_0$, $\hat{\theta}_0 = \hat{\theta}_0(a, p, n, \mu)$ and $\gamma = \gamma(a, p, n, \mu)$ such that, if $\theta \geq \hat{\theta}_0$, then $\mathcal{J}_a$ satisfies the following inequality for all $s_2 > s_1 \geq \max\{\hat{s}_0, -\log T\}$*:

$$\mathcal{J}_a[w](s_2) - \mathcal{J}_a[w](s_1) \leq -\frac{1}{2}\int_{s_1}^{s_2}\int_{\mathbb{R}^n}(\partial_s w)^2\rho\, dy\, ds. \tag{1-18}$$

As in [Giga et al. 2004a; Nguyen 2015], the existence of the Lyapunov functional is a crucial step for deriving the blow-up rate (1-4) and then the blow-up limit. In particular, we have the following:

Theorem 1.2. *Let a, p, n and μ be fixed and let u be a blow-up solution of* (1-1) *with a blow-up time T.*

(i) *Blow-up rate: There exists $\hat{s}_1 = \hat{s}_1(a, p, n, \mu) \geq \hat{s}_0$ such that, for all $s \geq s' = \max\{\hat{s}_1, -\log T\}$,*

$$\|w_{b,T}(y, s)\|_{L^\infty(\mathbb{R}^n)} \leq C, \tag{1-19}$$

where $w_{b,T}$ is as defined in (1-10) *and C is a positive constant depending only on n, p, μ and a bound of $\|w_{b,T}(\hat{s}_0)\|_{L^\infty}$.*

(ii) Blow-up limit: *If $\hat{a}$ is a blow-up point, then*

$$\lim_{t\to T}(T-t)^{\frac{1}{p-1}}u(\hat{a}+y\sqrt{T-t},t)=\lim_{s\to+\infty} w_{\hat{a},T}(y,s)=\pm\kappa \tag{1-20}$$

holds in L^2_ρ (L^2_ρ is the weighted L^2 space associated with the weight ρ of (1-12)) *and also uniformly on each compact subset of $\mathbb{R}^n$.*

Remark 1.3. We will not give the proof of Theorem 1.2 because its proof follows from Theorem 1.1 as in [Nguyen 2015]. Hence, we only give the proof of Theorem 1.1 and refer the reader to [Nguyen 2015, Section 2] for the proofs of (1-19) and (1-20).

The next step consists in obtaining an additional term in the asymptotic expansion given in Theorem 1.2(ii). Given b a blow-up point of $u(x,t)$, and up to changing u_0 by $-u_0$ and h by $-h$, we may assume that $w_{b,T}\to\kappa$ in L^2_ρ as $s\to+\infty$. As in [Nguyen 2015], we linearize $w_{b,T}$ around ϕ, where ϕ is the positive solution of the ordinary differential equation associated to (1-11),

$$\phi'=-\frac{\phi}{p-1}+\phi^p+e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}\phi) \tag{1-21}$$

such that

$$\phi(s)\to\kappa \quad \text{as } s\to+\infty; \tag{1-22}$$

see [Nguyen 2015, Lemma A.3] for the existence of ϕ, and note that ϕ is unique. For the reader's convenience, we give in Lemma A.1 the expansion of ϕ as $s\to+\infty$.

Let us introduce $v_{b,T}=w_{b,T}-\phi(s)$; then $\|v_{b,T}(y,s)\|_{L^2_\rho}\to 0$ as $s\to+\infty$ and $v_{b,T}$ (or v for simplicity) satisfies the equation

$$\partial_s v=(\mathcal{L}+\omega(s))v+F(v)+H(v,s) \quad \text{for all } y\in\mathbb{R}^n,\ s\in[-\log T,+\infty),$$

where $\mathcal{L}=\Delta-\frac{1}{2}y\cdot\nabla+1$ and ω, F and H satisfy

$$\omega(s)=\mathcal{O}\left(\frac{1}{s^{a+1}}\right) \quad \text{and} \quad |F(v)|+|H(v,s)|=\mathcal{O}(|v|^2) \quad \text{as } s\to+\infty,$$

(see the beginning of Section 3 for the proper definitions of ω, F and G).

It is well known that the operator $\mathcal{L}$ is self-adjoint in $L^2_\rho(\mathbb{R}^n)$. Its spectrum is given by

$$\operatorname{spec}(\mathcal{L})=\left\{1-\tfrac{1}{2}m \mid m\in\mathbb{N}\right\},$$

and it consists of eigenvalues. The eigenfunctions of $\mathcal{L}$ are derived from Hermite polynomials:

For $n=1$, the eigenfunction corresponding to $1-\frac{1}{2}m$ is

$$h_m(y)=\sum_{k=0}^{[m/2]}\frac{m!}{k!(m-2k)!}(-1)^k y^{m-2k}, \tag{1-23}$$

For $n\geq 2$, we write the spectrum of $\mathcal{L}$ as

$$\operatorname{spec}(\mathcal{L})=\left\{1-\tfrac{1}{2}|m| \;\middle|\; |m|=m_1+\cdots+m_n,\ (m_1,\dots,m_n)\in\mathbb{N}^n\right\}.$$

For $m = (m_1, \dots, m_n) \in \mathbb{N}^n$, the eigenfunction corresponding to $1 - \frac{1}{2}|m|$ is

$$H_m(y) = h_{m_1}(y_1) \cdots h_{m_n}(y_n), \tag{1-24}$$

where h_m is as defined in (1-23).

We also denote $c_m = c_{m_1} c_{m_2} \cdots c_{m_n}$ and $y^m = y_1^{m_1} y_2^{m_2} \cdots y_n^{m_n}$ for any $m = (m_1, \dots, m_n) \in \mathbb{N}^n$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$.

In this way, we derive the following asymptotic behaviors of $w_{b,T}(y, s)$ as $s \to +\infty$:

Theorem 1.4 (classification of the behavior of $w_{b,T}$ as $s \to +\infty$). *Consider a solution $u(t)$ of* (1-1) *which blows-up at time T and b a blow-up point. Let $w_{b,T}(y, s)$ be a solution of* (1-11). *Then one of the following possibilities occurs*:

(i) $w_{b,T}(y, s) \equiv \phi(s)$.

(ii) *There exists $l \in \{1, \dots, n\}$ such that, up to an orthogonal transformation of coordinates, we have*

$$w_{b,T}(y, s) = \phi(s) - \frac{\kappa}{4ps}\left(\sum_{j=1}^{l} y_j^2 - 2l\right) + \mathcal{O}\left(\frac{1}{s^{a+1}}\right) + \mathcal{O}\left(\frac{\log s}{s^2}\right) \quad \textit{as } s \to +\infty.$$

(iii) *There exist an integer $m \ge 3$ and constants c_α not all zero such that*

$$w_{b,T}(y, s) = \phi(s) - e^{-(\frac{m}{2}-1)s} \sum_{|\alpha|=m} c_\alpha H_\alpha(y) + o(e^{-(\frac{m}{2}-1)s}) \quad \textit{as } s \to +\infty.$$

The convergence takes place in L^2_ρ as well as in $\mathscr{C}^{k,\gamma}_{\text{loc}}$ for any $k \ge 1$ and some $\gamma \in (0, 1)$.

Remark 1.5. In [Nguyen 2015], we were unable to get this result in the case where h satisfies (1-9) with $a \in (0, 1]$. Here, by taking the particular form of the perturbation (see (1-2)), we are able to overcome technical difficulties in order to derive the result.

Remark 1.6. From Theorem 1.2(ii), we would naturally try to find an equivalent for $w - \kappa$ as $s \to +\infty$. A posteriori from our results in Theorem 1.4, we see that, in all cases, $\|w - \kappa\|_{L^2_\rho} \sim C/s^{a'}$ with $a' = \min\{a, 1\}$. This is indeed a new phenomenon observed in our (1-1) and which is different from the case of the unperturbed semilinear heat equation, where either $w - \kappa \equiv 0$ or $\|w - \kappa\|_{L^2_\rho} \sim C/s$ or $\|w - \kappa\|_{L^2_\rho} \sim Ce^{(1-m/2)s}$ for some even $m \ge 4$. This shows the originality of our paper. In our case, linearizing around κ would keep us trapped in the $1/s$ scale. In order to escape that scale, we forget the explicit function κ, which is not a solution of Equation (1-11) and linearize instead around the nonexplicit function ϕ, which happens to be an exact solution of (1-11). This way, we escape the $1/s$ scale and reach exponentially decreasing order.

Using the information obtained in Theorem 1.4, we can extend the asymptotic behavior of $w_{b,T}$ to larger regions. Particularly, we have the following:

Theorem 1.7 (convergence extension of $w_{b,T}$ to larger regions). *For all* $K_0 > 0$:

(i) *If Theorem 1.4(ii) occurs, then*

$$\sup_{|\xi|\le K_0} |w_{b,T}(\xi\sqrt{s}, s) - f_l(\xi)| = \mathcal{O}\left(\frac{1}{s^a}\right) + \mathcal{O}\left(\frac{\log s}{s}\right) \quad \text{as } s \to +\infty, \tag{1-25}$$

where

$$f_l(\xi) = \kappa\left(1 + \frac{p-1}{4p}\sum_{j=1}^{l} \xi_j^2\right)^{-\frac{1}{p-1}} \quad \text{for all } \xi \in \mathbb{R}^n \tag{1-26}$$

with l given in Theorem 1.4(ii).

(ii) *If Theorem 1.4(iii) occurs, then $m \ge 4$ is even and*

$$\sup_{|\xi|\le K_0} |w_{b,T}(\xi e^{(\frac{1}{2}-\frac{1}{m})s}, s) - \psi_m(\xi)| \to 0 \quad \text{as } s \to +\infty, \tag{1-27}$$

where

$$\psi_m(\xi) = \kappa\left(1 + \kappa^{-p}\sum_{|\alpha|=m} c_\alpha \xi^\alpha\right)^{-\frac{1}{p-1}} \quad \text{for all } \xi \in \mathbb{R}^n, \tag{1-28}$$

where c_α is the same as in Theorem 1.4 and the multilinear for $\sum_{|\alpha|=m} c_\alpha \xi^\alpha$ is nonnegative.

Remark 1.8. Note that Theorem 1.7 is analogous to the result obtained in [Velázquez 1992] for problem (1-1) without the perturbation. In particular, we follow the method of [loc. cit.] and care about the speed of the convergence, which was not given in that paper. Note also that the asymptotic profiles described in Theorem 1.7 are exactly the same as the ones derived in [loc. cit.] because we derived in this theorem the first-order approximation for the solution, unlike in Theorem 1.4, where we find the following terms in the expansion of the solution up to the second order. As in the unperturbed case ($h \equiv 0$), we expect that (1-25) is stable (see the previous remarks, particularly the paragraph after (1-5)) and (1-27) should correspond to unstable behaviors. The instability of (1-27) was proved only in one space dimension by Herrero and Velázquez [1992b; 1992d]. In particular, they proved the genericity of the asymptotic profile (1-25) in the one-dimensional case and announced the same for higher-dimensional cases, but they have never published it. While discussing numerical simulation for Equation (1-1) in one space dimension (see Section 4B below), we see that the numerical solutions exhibit only the behavior (1-25) and we could never obtain the behavior (1-27). This is probably due to the fact that the behavior (1-27) is unstable.

Remark 1.9. In [Nguyen and Zaag 2014], we constructed for the problem (1-1) with h given by (1-2) or (1-9) a solution which blows up in finite time at only one point and verifies the behavior (1-25) with $l = n$. The construction is inspired by the method of [Bricmont and Kupiainen 1994; Merle and Zaag 1997], relying on the reduction of the problem to a finite-dimensional one and a topological argument based on index theory.

At the end of this work, we give numerical confirmations for the asymptotic profile described in Theorem 1.7. For this purpose, we propose a new mesh-refinement method inspired by the rescaling algorithm of [Berger and Kohn 1988]. Note that their method was successful to solve blowing-up problems which are invariant under the transformation

$$\gamma \mapsto u_\gamma(\xi,\tau) = \gamma^{\frac{2}{p-1}} u(\gamma\xi, \gamma^2\tau) \quad \text{for all } \gamma > 0. \tag{1-29}$$

However, there are a lot of equations whose solutions blow up in finite time but which do not satisfy the property (1-29); one of them is (1-1) because of the presence of the perturbation term h. Although our method is very similar to Berger and Kohn's algorithm in spirit, it is better in the sense that it can be applied to a larger class of blowing-up problems which do not satisfy the rescaling property (1-29). To our knowledge, there are not many papers on the numerical blow-up profile, apart from [Berger and Kohn 1988] (see also [Nguyen 2014]), who already obtained numerical results for (1-1) without the perturbation term. For other numerical aspects, there are several studies for (1-1) in the unperturbed case; see, for example, [Abia, López-Marcos and Martínez 1998; 2001; Groisman and Rossi 2001; 2004; Groisman 2006; N'gohisse and Boni 2011; Kyza and Makridakis 2011; Cangiani et al. ≥ 2016] and the references therein. There is also the work of Baruch et al. [2010] studying standing-ring solutions.

This paper is organized as follows: Section 2 is devoted to the proof of Theorem 1.1. Theorem 1.2 follows from Theorem 1.1. Since all the arguments presented in [Nguyen 2015] remain valid for the case (1-9), except the existence of the Lyapunov functional for (1-11) (Theorem 1.1), we kindly refer the reader to [Nguyen 2015, Sections 2.3 and 2.4] for details of the proof. Section 3 deals with results on asymptotic behaviors (Theorems 1.4 and 1.7). In Section 4, we describe the new mesh-refinement method and give some numerical justifications for the theoretical results.

2. Existence of a Lyapunov functional for (1-11)

In this section, we mainly aim at proving that the functional $\mathcal{J}_a$ defined in (1-14) is a Lyapunov functional for (1-11) (Theorem 1.1). Note that this functional is far from being trivial and makes our main contribution.

In what follows, we denote by C a generic constant depending only on a, p, n and μ. We first give the following estimates on the perturbation term appearing in (1-11):

Lemma 2.1. *Let h be the function defined in* (1-2). *For all $\epsilon \in (0, p]$, there exists $C_0 = C_0(a,\mu,p,\epsilon) > 0$ and $\bar{s}_0 = \bar{s}_0(a,p,\epsilon) > 0$ large enough such that, for all $s \ge \bar{s}_0$,*

(i)
$$|e^{-\frac{ps}{p-1}} h(e^{\frac{s}{p-1}} z)| \le \frac{C_0}{s^a}(|z|^p + |z|^{p-\epsilon}),$$
$$|e^{-\frac{(p+1)s}{p-1}} H(e^{\frac{s}{p-1}} z)| \le \frac{C_0}{s^a}(|z|^{p+1} + 1),$$

where H is as defined in (1-17).

(ii)
$$|(p+1)e^{-\frac{(p+1)s}{p-1}} H(e^{\frac{s}{p-1}} z) - e^{-\frac{ps}{p-1}} h(e^{\frac{s}{p-1}} z) z| \le \tfrac{C_0}{s^{a+1}}(|z|^{p+1} + 1).$$

Proof. Note that (i) obviously follows from the estimate

$$\forall q > 0, \ \ \forall b > 0, \quad \frac{|z|^q}{\log^b(2 + e^{\frac{2s}{p-1}} z^2)} \le \frac{C}{s^b}(|z|^q + 1) \quad \text{for all } s \ge \bar{s}_0, \tag{2-1}$$

where $C = C(b, q) > 0$ and $\bar{s}_0 = \bar{s}_0(b, q) > 0$.

In order to derive the estimate (2-1), by considering the first case $z^2 e^{\frac{s}{p-1}} \ge 4$ then the case $z^2 e^{\frac{s}{p-1}} \le 4$, we would obtain (2-1).

Part (ii) directly follows from an integration by parts and the estimate (2-1). Indeed, we have

$$H(\xi) = \int_0^\xi h(x)\,dx = \mu \int_0^\xi \frac{|x|^{p-1}x}{\log^a(2+x^2)}\,dx$$
$$= \frac{\mu|\xi|^{p+1}}{(p+1)\log^a(2+\xi^2)} + \frac{2a\mu}{p+1}\int_0^\xi \frac{|x|^{p+1}x}{(2+x^2)\log^{a+1}(2+x^2)}\,dx.$$

Replacing ξ by $e^s/(p-1)z$ and using (2-1), we then derive (ii). This ends the proof of Lemma 2.1. □

We assert that Theorem 1.1 is a direct consequence of the following lemma:

Lemma 2.2. *Let a, p, n and μ be fixed and w be a solution of* (1-11). *There exists $\tilde{s}_0 = \tilde{s}_0(a, p, n, \mu) \ge s_0$ such that the functional of $\mathscr{E}$ defined in* (1-15) *satisfies the following inequality, for all $s \ge \max\{\tilde{s}_0, -\log T\}$:*

$$\frac{d}{ds}\mathscr{E}[w](s) \le -\frac{1}{2}\int_{\mathbb{R}^n} w_s^2 \rho\,dy + \gamma s^{-(a+1)}\mathscr{E}[w](s) + Cs^{-(a+1)}, \tag{2-2}$$

where $\gamma = 4C_0(p+1)/(p-1)^2$ and C_0 is given in Lemma 2.1.

Let us first derive Theorem 1.1 from Lemma 2.2, which we will prove later.

Proof of Theorem 1.1, given Lemma 2.2. Differentiating the functional $\mathscr{J}$ defined in (1-14), we obtain

$$\begin{aligned}\frac{d}{ds}\mathscr{J}_a[w](s) &= \frac{d}{ds}\{\mathscr{E}[w](s)e^{\frac{\gamma}{a}s^{-a}} + \theta s^{-a}\}\\ &= \frac{d}{ds}\mathscr{E}[w](s)e^{\frac{\gamma}{a}s^{-a}} - \gamma s^{-(a+1)}\mathscr{E}[w](s)e^{\frac{\gamma}{a}s^{-a}} - a\theta s^{-(a+1)}\\ &\le -\frac{1}{2}e^{\frac{\gamma}{a}s^{-a}}\int_{\mathbb{R}^n} w_s^2\rho\,dy + [Ce^{\frac{\gamma}{a}s^{-a}} - a\theta]s^{-(a+1)} \quad \text{(using (2-2))}.\end{aligned}$$

Choosing θ large enough that $Ce^{\gamma\tilde{s}_0^{-a}/a} - a\theta \le 0$ and noticing that $e^{\gamma s^{-a}/a} \ge 1$ for all $s > 0$, we derive

$$\frac{d}{ds}\mathscr{J}_a[w](s) \le -\frac{1}{2}\int_{\mathbb{R}^n} w_s^2\rho\,dy \quad \text{for all } s \ge \tilde{s}_0.$$

This implies the inequality (1-18) and concludes the proof of Theorem 1.1, assuming that Lemma 2.2 holds. □

Proof of Lemma 2.2. Multiplying (1-11) by $w_s\rho$ and integrating by parts,

$$\int_{\mathbb{R}^n}|w_s|^2\rho = -\frac{d}{ds}\left\{\int_{\mathbb{R}^n}\left(\frac{1}{2}|\nabla w|^2 + \frac{1}{2(p-1)}|w|^2 - \frac{1}{p+1}|w|^{p+1}\right)\rho\,dy\right\} + e^{-\frac{ps}{p-1}}\int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}}w)w_s\rho\,dy.$$

For the last term of the above expression, we obtain

$$
\begin{aligned}
e^{-\frac{ps}{p-1}} & \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w_s \rho \, dy \\
& = e^{-\frac{(p+1)s}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w)\Big(e^{\frac{s}{p-1}} w_s + \frac{e^{\frac{s}{p-1}}}{p-1} w\Big)\rho \, dy - \frac{1}{p-1} e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy \\
& = e^{-\frac{p+1}{p-1}s} \frac{d}{ds} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy - \frac{1}{p-1} e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy.
\end{aligned}
$$

This yields

$$
\begin{aligned}
\int_{\mathbb{R}^n} |w_s|^2 \rho \, dy = & -\frac{d}{ds}\Big\{\int_{\mathbb{R}^n}\Big(\frac{1}{2}|\nabla w|^2 + \frac{1}{2(p-1)}|w|^2 - \frac{1}{p+1}|w|^{p+1}\Big)\rho \, dy\Big\} \\
& + \frac{d}{ds}\Big\{e^{-\frac{p+1}{p-1}s} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy\Big\} \\
& + \frac{p+1}{p-1} e^{-\frac{p+1}{p-1}s} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy \\
& - \frac{1}{p-1} e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy.
\end{aligned}
$$

From the definition of the functional $\mathscr{E}$ given in (1-15), we derive a first identity in the following:

$$
\begin{aligned}
\frac{d}{ds} & \mathscr{E}[w](s) \\
& = -\int_{\mathbb{R}^n} |w_s|^2 \rho \, dy + \frac{p+1}{p-1} e^{-\frac{p+1}{p-1}s} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy - \frac{1}{p-1} e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy. \quad \text{(2-3)}
\end{aligned}
$$

A second identity is obtained by multiplying (1-11) by $w\rho$ and integrating by parts:

$$
\begin{aligned}
\frac{d}{ds} & \int_{\mathbb{R}^n} |w|^2 \rho \, dy \\
& = -4\Big\{\int_{\mathbb{R}^n}\Big(\frac{1}{2}|\nabla w|^2 + \frac{1}{2(p-1)}|w|^2 - \frac{1}{p+1}|w|^{p+1}\Big)\rho \, dy - e^{-\frac{(p+1)s}{p-1}} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy\Big\} \\
& \quad + \Big(2 - \frac{4}{p+1}\Big) \int_{\mathbb{R}^n} |w|^{p+1} \rho \, dy - 4 e^{-\frac{p+1}{p-1}s} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy + 2 e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy.
\end{aligned}
$$

Using the definition of $\mathscr{E}$ given in (1-15) again, we rewrite the second identity as follows:

$$
\begin{aligned}
\frac{d}{ds} \int_{\mathbb{R}^n} |w|^2 \rho \, dy = & -4\mathscr{E}[w](s) + 2\frac{p-1}{p+1} \int_{\mathbb{R}^n} |w|^{p+1} \rho \, dy \\
& - 4 e^{-\frac{p+1}{p-1}s} \int_{\mathbb{R}^n} H(e^{\frac{s}{p-1}} w)\rho \, dy + 2 e^{-\frac{ps}{p-1}} \int_{\mathbb{R}^n} h(e^{\frac{s}{p-1}} w) w \rho \, dy. \quad \text{(2-4)}
\end{aligned}
$$

From (2-3), we estimate

$$
\frac{d}{ds}\mathscr{E}[w](s) \le -\int_{\mathbb{R}^n} |w_s|^2 \rho \, dy + \frac{1}{p-1} \int_{\mathbb{R}^n} \Big\{\Big|(p+1) e^{-\frac{(p+1)s}{p-1}} H(e^{\frac{s}{p-1}} w) - e^{-\frac{ps}{p-1}} h(e^{\frac{s}{p-1}} w) w\Big|\Big\}\rho \, dy.
$$

Using Lemma 2.1(ii), we have, for all $s \geq \bar{s}_0$,

$$\frac{d}{ds}\mathscr{E}[w](s) \leq -\int_{\mathbb{R}^n} |w_s|^2 \rho\, dy + \frac{C_0 s^{-(a+1)}}{p-1}\int_{\mathbb{R}^n} |w|^{p+1}\rho\, dy + Cs^{-(a+1)}. \tag{2-5}$$

On the other hand, by (2-4) we have

$$\begin{aligned}\int_{\mathbb{R}^n} |w|^{p+1}\rho\, dy \leq \frac{2(p+1)}{p-1}\mathscr{E}[w](s) + \frac{p+1}{p-1}\int_{\mathbb{R}^n} |w_s w|\rho\, dy\\ + \frac{2(p+1)}{p-1}\int_{\mathbb{R}^n}\big(|e^{-\frac{p+1}{p-1}s}H(e^{\frac{s}{p-1}}w)| + |e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}w)w|\big)\rho\, dy.\end{aligned}$$

Using Lemma 2.1(i) and the fact that $|w_s w| \leq \epsilon(|w_s|^2 + |w|^{p+1}) + C(\epsilon)$ for all $\epsilon > 0$, we obtain

$$\int_{\mathbb{R}^n} |w|^{p+1}\rho\, dy \leq \frac{2(p+1)}{p-1}\mathscr{E}[w](s) + \epsilon\int_{\mathbb{R}^n} |w_s|^2\rho\, dy + (\epsilon + Cs^{-a})\int_{\mathbb{R}^n} |w|^{p+1}\rho\, dy + C.$$

Taking $\epsilon = \frac{1}{4}$ and s_1 large enough that $Cs^{-a} \leq \frac{1}{4}$ for all $s \geq s_1$, we have

$$\int_{\mathbb{R}^n} |w|^{p+1}\rho\, dy \leq \frac{4(p+1)}{p-1}\mathscr{E}[w](s) + \frac{1}{2}\int_{\mathbb{R}^n} |w_s|^2\rho\, dy + C \quad \text{for all } s > s_1. \tag{2-6}$$

Substituting (2-6) into (2-5) yields (2-2) with $\tilde{s}_0 = \max\{\bar{s}_0, s_1\}$. This concludes the proof of Lemma 2.2 and Theorem 1.1 also. □

3. Blow-up behavior

This section is devoted to the proof of Theorems 1.4 and 1.7. Consider a blow-up point b and write w instead of $w_{b,T}$ for simplicity. From Theorem 1.2(ii) and up to changing the signs of w and h, we may assume that $\|w(y,s) - \kappa\|_{L^2_\rho} \to 0$ as $s \to +\infty$ uniformly on compact subsets of $\mathbb{R}^n$. As mentioned in the introduction, by setting $v(y,s) = w(y,s) - \phi(s)$ (ϕ is the positive solution of (1-21) such that $\phi(s) \to \kappa$ as $s \to +\infty$), we see that $\|v(y,s)\|_{L^2_\rho} \to 0$ as $s \to +\infty$ and v solves the equation

$$\partial_s v = (\mathscr{L} + \omega(s))v + F(v) + G(v,s) \quad \text{for all } y \in \mathbb{R}^n,\ s \in [-\log T, +\infty), \tag{3-1}$$

where $\mathscr{L} = \Delta - \frac{1}{2}y\cdot\nabla + 1$ and ω, F and G are given by

$$\begin{aligned}\omega(s) &= p(\phi^{p-1} - \kappa^{p-1}) + e^{-s}h'(e^{\frac{s}{p-1}}\phi),\\ F(v) &= |v+\phi|^{p-1}(v+\phi) - \phi^p - p\phi^{p-1}v,\\ G(v,s) &= e^{-\frac{ps}{p-1}}\big[h(e^{\frac{s}{p-1}}(v+\phi)) - h(e^{\frac{s}{p-1}}\phi) - e^{\frac{s}{p-1}}h'(e^{\frac{s}{p-1}}\phi)v\big].\end{aligned}$$

By a direct calculation, we can show that

$$|\omega(s)| = \mathscr{O}\left(\frac{1}{s^{a+1}}\right) \quad \text{as } s \to +\infty \tag{3-2}$$

(see Lemma B.1 for the proof of this fact; note also that in the case where h is given by (1-9) and treated in [Nguyen 2015], we just obtain $|\omega(s)| = \mathscr{O}(s^{-a})$ as $s \to +\infty$, which was a major reason preventing us from deriving the result in the case $a \in (0,1]$) there.

Now, introducing

$$V(y,s)=\beta(s)v(y,s),\quad \text{where}\quad \beta(s)=\exp\left(-\int_s^{+\infty}\omega(\tau)\,d\tau\right), \tag{3-3}$$

V satisfies

$$\partial_s V=\mathscr{L}V+\bar{F}(V,s), \tag{3-4}$$

where $\bar{F}(V,s)=\beta(s)(F(V)+G(V,s))$ satisfies

$$\left|\bar{F}(V,s)-\frac{p}{2\kappa}V^2\right|=\mathbb{O}\left(\frac{V^2}{s^a}\right)+\mathbb{O}(|V|^3)\quad \text{as } s\to+\infty \tag{3-5}$$

(see [Nguyen 2015, Lemma C.1] for the proof of this fact; note that, in the case where h is given by (1-9), the first term in the right-hand side of (3-5) is $\mathbb{O}(V^2/s^{a-1})$).

Since $\beta(s)\to 1$ as $s\to+\infty$, each equivalent for V is also an equivalent for v. Therefore, it suffices to study the asymptotic behavior of V as $s\to+\infty$. More precisely, we claim the following:

Proposition 3.1 (classification of the behavior of V as $s\to+\infty$). *One of the following possibilities occurs*:

(i) $V(y,s)\equiv 0$.

(ii) *There exists $l\in\{1,\dots,n\}$ such that, up to an orthogonal transformation of coordinates, we have*

$$V(y,s)=-\frac{\kappa}{4ps}\left(\sum_{j=1}^{l}y_j^2-2l\right)+\mathbb{O}\left(\frac{1}{s^{a+1}}\right)+\mathbb{O}\left(\frac{\log s}{s^2}\right)\quad \textit{as } s\to+\infty.$$

(iii) *There exist an integer $m\geq 3$ and constants c_α not all zero such that*

$$V(y,s)=-e^{(1-\frac{m}{2})s}\sum_{|\alpha|=m}c_\alpha H_\alpha(y)+o(e^{(1-\frac{m}{2})s})\quad \textit{as } s\to+\infty.$$

The convergence takes place in L^2_ρ as well as in $\mathscr{C}^{k,\gamma}_{\text{loc}}$ for any $k\geq 1$ and $\gamma\in(0,1)$.

Proof. Because we have the same equation (3-4) and a similar estimate (3-5) to the case treated in [Nguyen 2015], we do not give the proof and kindly refer the reader to Section 3 there. ☐

Let us derive Theorem 1.4 from Proposition 3.1.

Proof of Theorem 1.4. By the definition (3-3) of V, we see that given Proposition 3.1(i) it directly follows that $v(y,s)\equiv\phi(s)$, which is Theorem 1.4(i). Using Proposition 3.1(ii) and the fact that $\beta(s)=1+\mathbb{O}(1/s^a)$ as $s\to+\infty$, we see that, as $s\to+\infty$,

$$w(y,s)=\phi(s)+V(y,s)\left(1+\mathbb{O}\left(\frac{1}{s^a}\right)\right)=\phi(s)-\frac{\kappa}{4ps}\left(\sum_{j=1}^{l}y_j^2-2l\right)+\mathbb{O}\left(\frac{1}{s^{a+1}}\right)+\mathbb{O}\left(\frac{\log s}{s^2}\right),$$

which yields Theorem 1.4(ii).

Using Proposition 3.1(iii) and again the fact that $\beta(s) = 1 + \mathcal{O}(1/s^a)$ as $s \to +\infty$, we have

$$w(y,s) = \phi(s) - e^{(1-\frac{m}{2})s} \sum_{|\alpha|=m} c_\alpha H_\alpha(y) + o(e^{(1-\frac{m}{2})s}) \quad \text{as } s \to +\infty.$$

This concludes the proof of Theorem 1.4. □

We now give the proof of Theorem 1.7 from Theorem 1.4. Note that the derivation of Theorem 1.7 from Theorem 1.4 in the unperturbed case ($h \equiv 0$) was done by Velázquez [1992]. The idea to extend the convergence up to sets of the type $\{|y| \le K_0\sqrt{s}\}$ or $\{|y| \le K_0 e^{(1/2-1/m)s}\}$ is to estimate the effect of the convective term $-\frac{1}{2} y \cdot \nabla w$ in (1-11) in L^q_ρ spaces with $q > 1$. Since the proof of Theorem 1.7 is, in spirit, by the method given in [Velázquez 1992], all that we need to do is to control the strong perturbation term in (1-11). We therefore give the main steps of the proof and focus only on the new arguments. Note also that we only give the proof of Theorem 1.4(ii) because the proof of (iii) is exactly the same as in Proposition 34 in [Nguyen 2015].

Let us restate Theorem 1.7(i) in the following proposition:

Proposition 3.2 (asymptotic behavior in the $y/\sqrt{s}$ variable). *Assume that w is a solution of* (1-11) *which satisfies Theorem 1.4(ii). Then, for all* $K > 0$,

$$\sup_{|\xi| \le K} |w(\xi\sqrt{s}, s) - f_l(\xi)| = \mathcal{O}\left(\frac{1}{s^a}\right) + \mathcal{O}\left(\frac{\log s}{s}\right) \quad \textit{as } s \to +\infty,$$

where $f_l(\xi) = \kappa\big(1 + ((p-1)/4p)\sum_{j=1}^{l} \xi_j^2\big)^{-1/(p-1)}$.

Proof. Define $q = w - \varphi$, where

$$\varphi(y,s) = \frac{\phi(s)}{\kappa}\left[\kappa\left(1 + \frac{p-1}{4ps}\sum_{j=1}^{l} y_j^2\right)^{-\frac{1}{p-1}} + \frac{\kappa l}{2ps}\right], \tag{3-6}$$

and ϕ is the unique positive solution of (1-21) satisfying (1-22).

Note that in [Velázquez 1992; Nguyen 2015], the authors took

$$\varphi(y,s) = \kappa\left(1 + \frac{p-1}{4ps}\sum_{j=1}^{l} y_j^2\right)^{-\frac{1}{p-1}} + \frac{\kappa l}{2ps}.$$

But this choice just works in the case where $a > 1$. In the particular case (1-2), we use in addition the factor $\phi(s)/\kappa$, which allows us to go beyond the order $1/s^a$ coming from the strong perturbation term in order to reach $1/s^{a+1}$ in many estimates in the proof.

Using Taylor's formula in (3-6) and Theorem 1.4(ii), we find that

$$\|q(y,s)\|_{L^2_\rho} = \mathcal{O}\left(\frac{1}{s^{a+1}}\right) + \mathcal{O}\left(\frac{\log s}{s^2}\right) \quad \text{as } s \to +\infty. \tag{3-7}$$

Straightforward calculations based on (1-11) yield

$$\partial_s q = (\mathcal{L} + \alpha) q + F(q) + G(q,s) + R(y,s) \quad \text{for all } (y,s) \in \mathbb{R}^n \times [-\log T, +\infty), \tag{3-8}$$

where

$$\begin{aligned}\alpha(y,s) &= p(\varphi^{p-1}-\kappa^{p-1})+e^{-s}h'(e^{\frac{s}{p-1}}\varphi),\\ F(q) &= |q+\varphi|^{p-1}(q+\varphi)-\varphi^p-p\varphi^{p-1}q,\\ G(q,s) &= e^{-\frac{ps}{p-1}}[h(e^{\frac{s}{p-1}}(q+\varphi))-h(e^{\frac{s}{p-1}}\varphi)-e^{\frac{s}{p-1}}h'(e^{\frac{s}{p-1}}\varphi)q],\\ R(y,s) &= -\partial_s\varphi+\Delta\varphi-\frac{y}{2}\cdot\nabla\varphi-\frac{\varphi}{p-1}+\varphi^p+e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}\varphi).\end{aligned}$$

Let $K_0>0$ be fixed; we consider first the case $|y|\geq 2K_0\sqrt{s}$ and then $|y|\leq 2K_0\sqrt{s}$ and make a Taylor expansion for $\xi=y/\sqrt{s}$ bounded. Simultaneously we obtain, for all $s\geq s_0$,

$$\begin{aligned}\alpha(y,s) &\leq \frac{C_1}{s^{a'}},\\ |F(q)|+|G(q,s)| &\leq C_1(q^2+\mathbf{1}_{\{|y|\geq 2K_0\sqrt{s}\}}),\\ |R(y,s)| &\leq C_1\left(\frac{|y|^2+1}{s^{1+a'}}+\mathbf{1}_{\{|y|\geq 2K_0\sqrt{s}\}}\right),\end{aligned}$$

where $a'=\min\{1,a\}$, $C_1=C_1(M_0,K_0)>0$ and M_0 is the bound of w in L^∞ norm. Note that we need to use in addition the fact that ϕ satisfies (1-21) to derive the bound for R (see Lemma B.2).

Let $Q=|q|$; we then use the above estimates and Kato's inequality, i.e., $\Delta f\cdot\operatorname{sign}(f)\leq\Delta(|f|)$, to derive from (3-8) the following: for all $K_0>0$ fixed, there are $C_*=C_*(K_0,M_0)>0$ and a time $s'>0$ large enough such that, for all $s\geq s_*=\max\{s',-\log T\}$,

$$\partial_s Q\leq\left(\mathscr{L}+\frac{C_*}{s^{a'}}\right)Q+C_*\left(Q^2+\frac{|y|^2+1}{s^{1+a'}}+\mathbf{1}_{\{|y|\geq 2K_0\sqrt{s}\}}\right)\quad\text{for all } y\in\mathbb{R}^n. \tag{3-9}$$

Since

$$\left|w(y,s)-f_l\left(\frac{y}{\sqrt{s}}\right)\right|\leq Q+\frac{C}{s^{a'}},$$

the conclusion of Proposition 3.2 follows if we show that

$$\forall K_0>0\qquad \sup_{|y|\leq K_0\sqrt{s}} Q(y,s)\to 0\quad\text{as } s\to+\infty. \tag{3-10}$$

Let us now focus on the proof of (3-10) in order to conclude Proposition 3.2. For this purpose, we introduce the following norm: for $r\geq 0$, $q>1$ and $f\in L^q_{\mathrm{loc}}(\mathbb{R}^n)$,

$$L_\rho^{q,r}(f)\equiv\sup_{|\xi|\leq r}\left(\int_{\mathbb{R}^n}|f(y)|^q\rho(y-\xi)\,dy\right)^{\frac{1}{q}}.$$

Following the idea in [Velázquez 1992], we shall make estimates on solutions of (3-9) in the $L_\rho^{2,r(\tau)}$ norm, where $r(\tau)=K_0e^{(\tau-\bar{s})/2}\leq K_0\sqrt{\tau}$. In particular, we have the following:

Lemma 3.3. *Let s be large enough and let $\bar{s}$ be defined by $e^{s-\bar{s}}=s$. Then, for all $\tau\in[\bar{s},s]$ and $K_0>0$,*

$$g(\tau)\leq C_0\left(e^{\tau-\bar{s}}\epsilon(\bar{s})+\int_{\bar{s}}^{(\tau-2K_0)_+}\frac{e^{\tau-t-2K_0}g^2(t)}{(1-e^{-(\tau-t-2K_0)})^{1/20}}\,dt\right),$$

where $g(\tau) = L_\rho^{2,r(K_0,\tau,\bar{s})}(Q(\tau))$, $r(K_0,\tau,\bar{s}) = K_0 e^{(\tau-\bar{s})/2}$, $\epsilon(s) = \mathbb{O}(1/s^{a+1}) + \mathbb{O}(\log s/s^2)$, $C_0 = C_0(C_*, M_0, K_0)$ *and* $z_+ = \max\{z, 0\}$.

Proof. Multiplying (3-9) by $\beta(\tau) = e^{\int_{\bar{s}}^{\tau} C_*/t^{a'}\,dt}$, we write $Q(y,\tau)$, for all $(y,\tau) \in \mathbb{R}^n \times [\bar{s}, s]$, in the integral form

$$Q(y,\tau) = \beta(\tau) S_{\mathscr{L}}(\tau-\bar{s}) Q(y,\bar{s}) + C_* \int_{\bar{s}}^{\tau} \beta(\tau) S_{\mathscr{L}}(\tau-t) \left(Q^2 + \frac{|y|^2}{t^{1+a'}} + \frac{1}{t^{1+a'}} + \mathbf{1}_{\{|y|\geq 2K_0\sqrt{t}\}} \right) dt,$$

where $S_{\mathscr{L}}$ is the linear semigroup corresponding to the operator $\mathscr{L}$.

Next, we take the $L_\rho^{2,r(K_0,\tau,\bar{s})}$ norms on both sides in order to get

$$\begin{aligned} g(\tau) &\leq C_0 L_\rho^{2,r}[S_{\mathscr{L}}(\tau-\bar{s}) Q(\bar{s})] + C_0 \int_{\bar{s}}^{\tau} L_\rho^{2,r}[S_{\mathscr{L}}(\tau-t) Q^2(t)]\,dt \\ &\quad + C_0 \int_{\bar{s}}^{\tau} L_\rho^{2,r}\left[S_{\mathscr{L}}(\tau-t) \left(\frac{|y|^2}{t^{1+a'}} + \frac{1}{t^{1+a'}} \right) \right] dt \\ &\quad + C_0 \int_{\bar{s}}^{\tau} L_\rho^{2,r}[S_{\mathscr{L}}(\tau-t) \mathbf{1}_{\{|y|\geq 2K_0\sqrt{t}\}}]\,dt \\ &\equiv J_1 + J_2 + J_3 + J_4. \end{aligned}$$

Proposition 2.3 in [Velázquez 1992] yields

$$\begin{aligned} |J_1| &\leq C_0 e^{\tau-\bar{s}} \|Q(\bar{s})\|_{L_\rho^2} = e^{\tau-\bar{s}} \mathbb{O}(\epsilon(\bar{s})) \quad \text{as } \bar{s} \to +\infty, \\ |J_2| &\leq \frac{C_0}{\bar{s}^{1+a'}} e^{\tau-\bar{s}} + C_0 \int_{\bar{s}}^{(\tau-2K_0)_+} \frac{e^{(\tau-t-2K_0)}}{(1-e^{-(\tau-t-2K_0)})^{1/20}} [L_\rho^{2,r(K_0,t,\bar{s})} Q(t)]^2\,dt, \\ |J_3| &\leq \frac{C_0 e^{\tau-\bar{s}}}{\bar{s}^{1+a'}} (1 + (\tau-\bar{s})), \\ |J_4| &\leq C_0 e^{-\delta\bar{s}}, \quad \text{where} \quad \delta = \delta(K_0) > 0. \end{aligned}$$

Putting together the estimates on J_i, $i = 1, 2, 3, 4$, we conclude the proof of Lemma 3.3. ☐

We now use the following Gronwall lemma:

Lemma 3.4 [Velázquez 1992]. *Let* ϵ, C, R *and* δ *be positive constants with* $\delta \in (0,1)$. *Assume that* $\mathscr{H}(\tau)$ *is a family of continuous functions satisfying*

$$\mathscr{H}(\tau) \leq \epsilon e^{\tau} + C \int_0^{(\tau-R)_+} \frac{e^{\tau-s} \mathscr{H}^2(s)}{(1-e^{-(\tau-s-R)})^{\delta}}\,ds \quad \text{for } \tau > 0.$$

Then there exist $\theta = \theta(\delta, C, R)$ *and* $\epsilon_0 = \epsilon_0(\delta, C, R)$ *such that, for all* $\epsilon \in (0, \epsilon_0)$ *and any* τ *for which* $\epsilon e^{\tau} \leq \theta$, *we have*

$$\mathscr{H}(\tau) \leq 2\epsilon e^{\tau}.$$

Applying Lemma 3.4 with $\mathscr{H} \equiv g$, we see from Lemma 3.3 that, for s large enough,

$$g(\tau) \leq 2C_0 e^{\tau-\bar{s}} \epsilon(\bar{s}) \quad \text{for all } \tau \in [\bar{s}, s].$$

If $\tau = s$, then $e^{s-\bar{s}} = s$, $r = K_0\sqrt{s}$ and

$$g(s) \equiv L_\rho^{2,K_0\sqrt{s}}(Q(s)) = \mathbb{O}\left(\frac{1}{s^a}\right) + \mathbb{O}\left(\frac{\log s}{s}\right) \quad \text{as } s \to +\infty.$$

By using the regularizing effects of the semigroup $S_{\mathcal{L}}$ (see [Velázquez 1992, Proposition 2.3]), we then obtain

$$\sup_{|y|\le K_0\sqrt{s}/2} Q(y,s) \le C'(C_*, K_0, M_0) L_\rho^{2,K_0\sqrt{s}}(Q(s)) = \mathbb{O}\left(\frac{1}{s^a}\right) + \mathbb{O}\left(\frac{\log s}{s}\right) \quad \text{as } s \to +\infty,$$

which concludes the proof of Proposition 3.2. □

4. Numerical method

We give in this section a numerical study of the blow-up profile of (1-1) in one dimension. Though our method is very similar to Berger and Kohn's algorithm [1988] in spirit, it is better in the sense that it can be applied to equations which are not invariant under the transformation (1-29). Our method differs from Berger and Kohn's in the following way: we step the solution forward until its maximum value multiplied by a power of its mesh size reaches a preset threshold, where the mesh size and the preset threshold are linked; for the rescaling algorithm, the solution is stepped forward until its maximum value reaches a preset threshold, and the mesh size and the preset threshold do not need to be linked. For more clarity, we present in the next subsection the mesh-refinement technique applied to (1-1), then give various numerical experiments to illustrate the effectiveness of our method for the problem of the numerical blow-up profile. Note that our method is more general than Berger and Kohn's, in the sense that it applies to non-scale-invariant equations. However, when applied to the unperturbed case $F(u) = |u|^{p-1}u$, our method gives exactly the same approximation as that of [Berger and Kohn 1988].

4A. *Mesh-refinement algorithm.* As usually with numerical simulations of blow-up (see [Berger and Kohn 1988]), we will simulate the equation on a bounded interval (say $(-A, A)$ with $A > 0$) with homogeneous Dirichlet boundary conditions, rather than the whole line $\mathbb{R}$. This choice is reasonable for two reasons:

- If initial data on the line are symmetric and decreasing to zero at infinity, then this property persists in time; hence, we are close to the situation of a bounded interval $(-A, A)$ with $A > 0$ large and homogeneous Dirichlet condition.
- We believe that the blow-up on a bounded interval is the same as on the whole line, given that blow-up does not occur on the boundary, as is already known for the pure power and $\mu = 0$. Moreover, as in [Giga and Kohn 1987; Giga et al. 2004b], the results stated in the introduction can be extended to the case when the problem (1-1) is considered in a convex domain of $\mathbb{R}^n$ with Dirichlet condition. Thus, they hold for the problem (4-1).

For that reason we focus on the bounded interval case $(-A, A)$ here. For simplicity we will take $A = 1$. In this section, we describe our refinement algorithm to solve numerically the problem (1-1) with initial data $\varphi(x) > 0$, $\varphi(x) = \varphi(-x)$, $x\, d\varphi(x)/dx < 0$ for $x \neq 0$, which gives a positive symmetric and radially decreasing solution. Let us rewrite the problem (1-1) (with $\mu = 1$) as

$$\begin{cases} \partial_t u = \partial_x^2 u + F(u), & (x,t) \in (-1,1) \times (0,T), \\ u(1,t) = u(-1,t) = 0, & t \in (0,T), \\ u(x,0) = \varphi(x), & x \in (-1,1), \end{cases} \tag{4-1}$$

where $p > 1$ and

$$F(u) = u^p + \frac{u^p}{\log^a(2+u^2)} \quad \text{with} \quad a > 0. \tag{4-2}$$

Let δ and τ be the initial space and time steps, we define $C_\Delta = \tau/\delta^2$, $x_i = i\delta$, $t_n = n\tau$, $I = 1/\delta$ and $u_{i,n}$ as the approximation of $u(x_i, t_n)$, where $u_{i,n}$ is defined for all $n \geq 0$ and $i \in \{-I, \dots, I\}$ by

$$\begin{aligned} u_{i,n+1} &= u_{i,n} + C_\Delta[u_{i-1,n} - 2u_{i,n} + u_{i+1,n}] + \tau F(u_{i,n}), \\ u_{I,n} &= u_{-I,n} = 0, \\ u_{i,0} &= \varphi_i. \end{aligned} \tag{4-3}$$

Note that this scheme is first-order accurate in time and second-order in space, and it requires the stability condition $C_\Delta = \tau/\delta^2 \leq \frac{1}{2}$.

Our algorithm needs to fix the following parameters:

- $\lambda < 1$, the refining factor with λ^{-1} being a small integer.
- M, the threshold to control the amplitude of the solution.
- α, the parameter controlling the width of interval to be refined.

The parameters λ and M must satisfy the relation

$$M = \lambda^{-\frac{2}{p-1}} M_0, \quad \text{where} \quad M_0 = \delta^{\frac{2}{p-1}} \|\varphi\|_\infty. \tag{4-4}$$

Note that the relation (4-4) is important to make our method work. In [Berger and Kohn 1988], the typical choice is $M_0 = \|\varphi\|_\infty$, hence $M = \lambda^{-2/(p-1)} \|\varphi\|_\infty$.

In the initial step of the algorithm, we simply apply the scheme (4-3) until $\delta^{2/(p-1)} \|u(\cdot, t_n)\|_\infty$ reaches M (note that in [Berger and Kohn 1988] the solution is stepped forward until $\|u(\cdot, t_n)\|_\infty$ reaches M; in this first step, the thresholds of the two methods are the same, however, they will split after the second step; roughly speaking, for the threshold we shall use the quantity $\delta^{2/(p-1)} \|u(\cdot, t_n)\|_\infty$ in our method instead of the $\|u(\cdot, t_n)\|_\infty$ in [Berger and Kohn 1988]). Then, we use a linear interpolation in time to find τ_0^* such that

$$t_n - \tau \leq \tau_0^* \leq t_n \quad \text{and} \quad \delta^{\frac{2}{p-1}} \|u(\cdot, \tau_0^*)\| = M.$$

Afterward, we determine two grid points y_0^- and y_0^+ such that

$$\begin{cases} \delta^{\frac{2}{p-1}} u(y_0^- - \delta, \tau_0^*) < \alpha M \leq \delta^{\frac{2}{p-1}} u(y_0^-, \tau_0^*), \\ \delta^{\frac{2}{p-1}} u(y_0^+ + \delta, \tau_0^*) < \alpha M \leq \delta^{\frac{2}{p-1}} u(y_0^+, \tau_0^*). \end{cases} \tag{4-5}$$

Note that $y_0^- = -y_0^+$ because of the symmetry of the solution. This finishes the initial step.

Let us begin the first refining step. Define

$$u^{(1)}(y^{(1)}, t^{(1)}) = u(y^{(1)}, \tau_0^* + t^{(1)}), \quad y^{(1)} \in (y_0^-, y_0^+),\ t^{(1)} \geq 0, \tag{4-6}$$

and set $\delta^{(1)} = \lambda\delta$, $\tau^{(1)} = \lambda^2\tau$ as the space and time step for the approximation of $u^{(1)}$ (note that $\tau^{(1)}/(\delta^{(1)})^2 = \tau/\delta^2 = C_\Delta$, which is a constant), $y_i^{(1)} = i\delta^{(1)}$, $t_n^{(1)} = n\tau^{(1)}$, $I_1 = y_0^+/\delta^{(1)}$ and $u_{i,n}^{(1)}$ as the approximation of $u^{(1)}(y_i^{(1)}, t_n^{(1)})$. Note that, in the unperturbed case, Berger and Kohn used the transformation (1-29) to define $u^{(1)}(y^{(1)}, t^{(1)}) = \lambda^{2/(p-1)}u(\lambda y^{(1)}, \tau_0^* + \lambda^2 t^{(1)})$ and then applied the same scheme for u to $u^{(1)}$. However, we can not do the same because (4-1) is not invariant under the transformation (1-29). Then applying the scheme (4-3) to $u^{(1)}$, we write

$$u_{i,n+1}^{(1)} = u_{i,n}^{(1)} + C_\Delta[u_{i-1,n}^{(1)} - 2u_{i,n}^{(1)} + u_{i+1,n}^{(1)}] + \tau^{(1)}F(u_{i,n}^{(1)}) \tag{4-7}$$

for all $n \geq 0$ and $i \in \{-I_1 + 1, \dots, I_1 - 1\}$.

Note that the computation of $u^{(1)}$ requires the initial data $u^{(1)}(y^{(1)}, 0)$ and the boundary condition $u^{(1)}(y_0^\pm, t^{(1)})$. For the initial condition, it is determined from $u(x, \tau_0^*)$ by using interpolation in space to get values at the new grid points. For the boundary condition, since $\tau^{(1)} = \lambda^2\tau$, we have from (4-6) that

$$u^{(1)}(y_0^\pm, n\tau^{(1)}) = u(y_0^\pm, \tau_0^* + n\lambda^2\tau). \tag{4-8}$$

Since u and $u^{(1)}$ will be stepped forward, each on its own grid ($u^{(1)}$ on (y_0^-, y_0^+) with the space and time steps $\delta^{(1)}$ and $\tau^{(1)}$, and u on $(-1, 1)$ with the space and time steps δ and τ), the relation (4-8) will provide us with the boundary values for $u^{(1)}$. In order to better understand how it works, let us consider an example with $\lambda = \frac{1}{2}$. After concluding the initial phase, the two solutions $u^{(1)}$ and u are stepped forward independently, each on its own grid; in other words, $u^{(1)}$ on (y_0^-, y_0^+) with the space and time steps $\delta^{(1)}$ and $\tau^{(1)}$, and u on $(-1, 1)$ with the space and time steps δ and τ. Then, using the linear interpolation in time for u, we get the boundary values for $u^{(1)}$ by (4-8), since $\tau^{(1)} = \lambda^2\tau = \frac{1}{4}\tau$. This means that u is stepped forward once every 4 time steps of $u^{(1)}$. After 4 steps forward of $u^{(1)}$, the values of u on the interval (y_0^-, y_0^+) must be updated to agree with the calculations of $u^{(1)}$. In other words, the approximation of u is used to assist in computing the boundary values for $u^{(1)}$. At each successive time step for u, the values of u on the interval (y_0^-, y_0^+) must be updated to make them agree with the more accurate fine grid solution $u^{(1)}$. When $(\delta^{(1)})^{2/(p-1)}\|u^{(1)}(\cdot, n\tau^{(1)})\|_\infty$ first exceeds M, we use a linear interpolation in time to find $\tau_1^* \in [\tau_{n-1}^{(1)}, \tau_n^{(1)}]$ such that $(\delta^{(1)})^{2/(p-1)}\|u^{(1)}(\cdot, \tau_1^*)\|_\infty = M$. On the interval where $(\delta^{(1)})^{2/(p-1)}\|u^{(1)}(\cdot, \tau_1^*)\|_\infty > \alpha M$, the grid is refined further and the entire procedure for $u^{(1)}$ is repeated to yield $u^{(2)}$, and so forth.

Before going to a general step, we would like to comment on relation (4-4). When $\delta^{2/(p-1)}\|u(\cdot, t)\|_\infty$ reaches the given threshold M in the initial phase, namely when $\delta^{2/(p-1)}\|u(\cdot, \tau_0^*)\|_\infty = M$, we want to refine the grid so that the maximum value of $(\delta^{(1)})^{2/(p-1)}u^{(1)}(y^{(1)}, 0)$ equals M_0. By (4-6), this request turns into $(\delta^{(1)})^{2/(p-1)}\|u(\cdot, \tau_0^*)\|_\infty = M_0$. Since $\delta^{(1)} = \lambda\delta$, it follows that $M = \lambda^{-2/(p-1)}M_0$, which yields (4-4).

Let $k \geq 0$; we set $\delta^{(k+1)} = \lambda^{-1}\delta^{(k)}$ and $\tau^{(k+1)} = \lambda^2\tau^{(k)}$ (note that $\tau^{(k+1)}/(\delta^{(k+1)})^2 = \tau^{(k)}/(\delta^{(k)})^2 = \cdots = \tau/\delta^2 = C_\Delta$), and let $y^{(k)}$ and $t^{(k)}$ be the variables of $u^{(k)}$, with $y_i^{(k)} = i\delta^{(k)}$ and $t_n^{(k)} = n\tau^{(k)}$. By convention, the index $k = 0$ means that $u^{(0)}(y^{(0)}, t^{(0)}) \equiv u(x,t)$, $\delta^{(0)} \equiv \delta$ and $\tau^{(0)} \equiv \tau$. The solution $u^{(k+1)}$ is related to $u^{(k)}$ by

$$u^{(k+1)}(y^{(k+1)}, t^{(k+1)}) = u^{(k)}(y^{(k+1)}, \tau_k^* + t^{(k+1)}), \tag{4-9}$$

where $y^{(k+1)} \in (y_k^-, y_k^+)$, $t^{(k+1)} \geq 0$, the time $\tau_k^* \in [t_{n-1}^{(k)}, t_n^{(k)}]$ satisfies

$$(\delta^{(k)})^{\frac{2}{p-1}} \|u^{(k)}(\cdot, \tau_k^*)\|_\infty = M,$$

and y_k^- and y_k^+ are two grid points determined by

$$\begin{cases} (\delta^{(k)})^{\frac{2}{p-1}} u^{(k)}(y_k^- - \delta^{(k)}, \tau_k^*) < \alpha M \leq (\delta^{(k)})^{\frac{2}{p-1}} u^{(k)}(y_k^-, \tau_k^*), \\ (\delta^{(k)})^{\frac{2}{p-1}} u^{(k)}(y_k^+ + \delta^{(k)}, \tau_k^*) < \alpha M \leq (\delta^{(k)})^{\frac{2}{p-1}} u^{(k)}(y_k^+, \tau_k^*). \end{cases} \tag{4-10}$$

The approximation of $u^{(k+1)}$ at the point $(y_i^{(k+1)}, t_n^{(k+1)})$, denoted by $u_{i,n}^{(k+1)}$, uses the scheme (4-3) with the space step $\delta^{(k+1)}$ and the time step $\tau^{(k+1)}$, which reads

$$u_{i,n+1}^{(k+1)} = u_{i,n}^{(k+1)} + C_\Delta[u_{i-1,n}^{(k+1)} - 2u_{i,n}^{(k+1)} + u_{i+1,n}^{(k+1)}] + \tau^{(k+1)} F(u_{i,n}^{(k+1)}) \tag{4-11}$$

for all $n \geq 1$ and $i \in \{-I_k + 1, \cdots, I_k - 1\}$, where $I_k = y_k^+/\delta^{(k+1)}$ (note that I_k is an integer since $\lambda^{-1} \in \mathbb{N}$).

As for the approximation of $u^{(k)}$, the computation of $u_{i,n}^{(k+1)}$ needs the initial data and the boundary condition. From (4-9) and the fact that $\tau^{(k+1)} = \lambda^2\tau^{(k)}$, we see that

$$u^{(k+1)}(y^{(k+1)}, 0) = u^{(k)}(y^{(k+1)}, \tau_k^*), \tag{4-12}$$

$$u^{(k+1)}(y_k^\pm, n\tau^{(k+1)}) = u^{(k)}(y_k^\pm, \tau_k^* + n\lambda^2\tau^{(k)}). \tag{4-13}$$

From (4-12), the initial data is simply calculated from $u^{(k)}(\cdot, \tau_k^*)$ by using a linear interpolation in space in order to assign values at new grid points. The essential step in this new mesh-refinement method is to determine the boundary condition through the identity (4-13), which means by a linear interpolation in time of $u^{(k)}$. Therefore, the previous solutions $u^{(k)}, u^{(k-1)}, \ldots$ are stepped forward independently, each on its own grid. More precisely, $\tau^{(k+1)} = \lambda^2\tau^{(k)} = \lambda^4\tau^{(k-1)} = \cdots$, so $u^{(k)}$ is stepped forward once every λ^{-2} time steps of $u^{(k+1)}$, $u^{(k-1)}$ once every λ^{-4} time steps of $u^{(k+1)}$, etc. On the other hand, the values of $u^{(k)}, u^{(k-1)}, \ldots$ must be updated to agree with the calculation of $u^{(k+1)}$. When $(\delta^{(k+1)})^{2/(p-1)} \|u^{(k+1)}(\cdot, \tau^{(k+1)})\|_\infty > M$, it is time for the next refining phase.

We would like to comment on the output of the refinement algorithm:

(i) Let τ_k^* be the time at which the refining takes place, then the ratio $\tau_k^*/\tau^{(k)}$, which indicates the number of time steps until $(\delta^{(k)})^{2/(p-1)}\|u^{(k)}\|_\infty$ reaches the given threshold M, tends to a constant as $k \to \infty$.

(ii) Let $u^{(k)}(\cdot, \tau_k^*)$ be the *refining solution*. If we plot $(\delta^{(k)})^{2/(p-1)} u^{(k)}(\cdot, \tau_k^*)$ on $(-1, 1)$, then their graphs eventually converge to a predicted one as $k \to \infty$.

(iii) Let (y_k^-, y_k^+) be the interval to be refined; then the quantity $(\delta^{(k)})^{-2}(y_k^+)^2$ behaves as a linear function of k.

These assertions can be well understood by the following theorem:

Theorem 4.1 (formal analysis). *Let u be a blowing-up solution to* (4-1); *then the output of the refinement algorithm satisfies*:

(i) *The ratio $\tau_k^*/\tau^{(k)}$ tends to a constant as $k \to \infty$, namely*

$$\frac{\tau_k^*}{\tau^{(k)}} \to \frac{(\lambda^{-2}-1)M^{1-p}}{C_\Delta(p-1)} \quad \text{as } k \to +\infty. \tag{4-14}$$

(ii) *Assume in addition that Theorem 1.7*(*i*) *holds. Defining $v^{(k)}(z) = (\delta^{(k)})^{2/(p-1)} u^{(k)}(zy_{k-1}^+, \tau_k^*)$ for all $k \geq 1$, we have*

$$\forall |z| < 1 \quad v^{(k)}(z) \sim M(1 + (\alpha^{1-p}-1)\lambda^{-2}z^2)^{-\frac{1}{p-1}} \quad \text{as } k \to +\infty. \tag{4-15}$$

(iii) *The quantity $(\delta^{(k)})^{-2}(y_k^+)^2$ behaves as a linear function, namely*

$$(\delta^{(k)})^{-2}(y_k^+)^2 \sim \gamma k + B \quad \text{as } k \to +\infty, \tag{4-16}$$

where

$$\gamma = \frac{2M^{1-p}(\alpha^{1-p}-1)|\log\lambda|}{c_p(p-1)\lambda^2}, \quad B = -\frac{M^{1-p}(\alpha^{1-p}-1)}{c_p(p-1)\lambda^2}\log\left(\frac{M^{1-p}\delta^2}{p-1}\right) \quad \text{and} \quad c_p = \frac{p-1}{4p}.$$

Remark 4.2. Note that there is no assumption on the value of a in the hypothesis in Theorem 4.1. It is understood in the sense that u blows up in finite time and its profile is described in Theorem 1.7.

Proof. As we will see in the proof, the statement (i) concerns the blow-up limit of the solution and (ii) is due to the blow-up profile stated in Theorem 1.7.

(i) If σ_k is the real time when the refinement from $u^{(k)}$ to $u^{(k+1)}$ takes place, we have, by (4-9),

$$\sigma_k = \tau_0^* + \tau_1^* + \cdots + \tau_k^*,$$

where τ_k^* is such that $(\delta^{(k)})^{2/(p-1)}\|u^{(k)}(\cdot\,, \tau_k^*)\|_\infty = M$. This means that

$$u^{(k)}(\cdot\,, \tau_k^*) = u(\cdot\,, \sigma_k). \tag{4-17}$$

On the other hand, from Theorem 1.7(i) and the definition (1-26) of f, we see that

$$\lim_{t\to T}(T-t)^{\frac{1}{p-1}}\|u(\cdot\,, t)\|_{L^\infty} = \kappa. \tag{4-18}$$

Combining (4-18) and (4-17) yields

$$(T-\sigma_k)^{\frac{1}{p-1}}\|u^{(k)}(\cdot\,, \tau_k^*)\|_\infty = \kappa + o(1), \tag{4-19}$$

where $o(1)$ represents a term that tends to 0 as $k \to +\infty$.

Since $\|u^{(k)}(\cdot,\tau_k^*)\|_\infty = M(\delta^{(k)})^{-2/(p-1)}$, we then derive

$$T-\sigma_k = (M^{-1}\kappa)^{p-1}(\delta^{(k)})^2 + o(1). \tag{4-20}$$

By the definition of σ_k and (4-17), we infer that $\tau_k^* = \sigma_k - \sigma_{k-1}$ (we can think τ_k^* as the *live time* of $u^{(k)}$ in the k-th refining phase). Hence,

$$\begin{aligned}\frac{\tau_k^*}{\tau^{(k)}} = \frac{\sigma_k-\sigma_{k-1}}{\tau^{(k)}} &= \frac{1}{\tau^{(k)}}[(T-\sigma_{k-1})-(T-\sigma_k)]\\ &= \frac{1}{\tau^{(k)}}(M^{-1}\kappa)^{p-1}((\delta^{(k-1)})^2-(\delta^{(k)})^2)+o(1)\\ &= \frac{(\delta^{(k)})^2}{\tau^{(k)}}(M^{-1}\kappa)^{p-1}(\lambda^{-2}-1)+o(1).\end{aligned}$$

Since the ratio $\tau^{(k)}/(\delta^{(k)})^2$ is always fixed by the constant C_Δ, we finally obtain

$$\lim_{k\to+\infty}\frac{\tau_k^*}{\tau^{(k)}} = \frac{(\lambda^{-2}-1)M^{1-p}}{C_\Delta(p-1)},$$

which is the conclusion of Theorem 4.1(i).

(ii) By the symmetry of the solution, we have $y_{k-1}^- = y_{k-1}^+$. We then consider the following mapping: for all $k\ge 1$,

$$z\mapsto v^{(k)}(z)\quad\text{for all }|z|\le 1,\quad\text{where}\quad v^{(k)}(z) = \big(\delta^{(k)}\big)^{\frac{2}{p-1}}u^{(k)}(zy_{k-1}^+,\tau_k^*).$$

We will show that $v^{(k)}(z)$ converges to some fixed function as $k\to+\infty$. For this purpose, we first write $u^{(k)}(y^{(k)},\tau_k*)$ in terms of $w(\xi,s)$ thanks to (4-17) and (1-10):

$$u^{(k)}(y^{(k)},\tau_k^*) = u(y^{(k)},\sigma_k) = (T-\sigma_k)^{-\frac{1}{p-1}}w(\xi^{(k)},s_k), \tag{4-21}$$

where $\xi^{(k)} = y^{(k)}/\sqrt{T-\sigma_k}$ and $s_k = -\log(T-\sigma_k)$.

If we write Theorem 1.7(i) in the variable $y/\sqrt{s}$ through (1-10), we have the equivalence

$$\left\| w(y,s) - f\left(\frac{y}{\sqrt{s}}\right)\right\|_{L^\infty} \to 0 \quad\text{as } s\to+\infty, \tag{4-22}$$

where f is as given in (1-26).

From (4-22), (4-20) and (4-21), we derive

$$u^{(k)}(y^{(k)},\tau_k^*) = M\kappa^{-1}(\delta^{(k)})^{-\frac{2}{p-1}} f\left(\frac{y^{(k)}}{(M^{-1}\kappa)^{\frac{p-1}{2}}\delta^{(k)}\sqrt{s_k}}\right)+o(1).$$

Then, multiplying both of sides by $(\delta^{(k)})^{2/(p-1)}$ and replacing $y^{(k)}$ by zy_{k-1}^+, we obtain

$$(\delta^{(k)})^{\frac{2}{p-1}}u^{(k)}(zy_{k-1}^+,\tau_k^*) = M\kappa^{-1} f\left(\frac{zy_{k-1}^+}{(M^{-1}\kappa)^{\frac{p-1}{2}}\delta^{(k)}\sqrt{s_k}}\right)+o(1). \tag{4-23}$$

From the definition (4-10) of y^+_{k-1}, we may assume that

$$(\delta^{(k-1)})^{\frac{2}{p-1}} u^{(k-1)}(y^+_{k-1}, \tau^*_{k-1}) = \alpha M.$$

Combining this with (4-23), we have

$$\alpha = \kappa^{-1} f\left(\frac{y^+_{k-1}}{(M^{-1}\kappa)^{\frac{p-1}{2}} \delta^{(k-1)} \sqrt{s_{k-1}}}\right) + o(1).$$

Since $s_k = -\log(T - \sigma_k)$ and $\delta^{(k)} = \lambda^k \delta$, we have from (4-20) that

$$s_k = 2k|\log\lambda| - \log\left(\frac{M^{1-p}\delta^2}{p-1}\right) + o(1), \tag{4-24}$$

which implies $\lim_{k\to+\infty} s_{k-1}/s_k = 1$. Thus, it is reasonable to assume that $y^+_{k-1}/\sqrt{s_{k-1}}$ and $y^+_{k-1}/\sqrt{s_k}$ tend to the positive root ζ as $k \to +\infty$. Hence,

$$\alpha = \kappa^{-1} f\left(\frac{\zeta}{(M^{-1}\kappa)^{\frac{p-1}{2}} \delta^{(k)} \lambda^{-1}}\right) + o(1).$$

Using the definition (1-26) of f, we have

$$\alpha = \left(1 + c_p \left|\frac{\zeta}{(M^{-1}\kappa)^{\frac{p-1}{2}} \delta^{(k)}}\right|^2 \lambda^2\right)^{-\frac{1}{p-1}} + o(1),$$

which implies

$$\left|\frac{\zeta}{(M^{-1}\kappa)^{\frac{p-1}{2}} \delta^{(k)}}\right|^2 = \frac{1}{c_p}[(\alpha^{1-p} - 1)\lambda^{-2}] + o(1), \tag{4-25}$$

where c_p is the constant given in the definition (1-26) of f.

Substituting this into (4-23) and using the definition (1-26) of f again, we arrive at

$$v^{(k)}(z) = M\left(1 + c_p \left|\frac{\zeta}{(M^{-1}\kappa)^{\frac{p-1}{2}} \delta^{(k)}}\right|^2 z^2\right)^{-\frac{1}{p-1}} + o(1) = M(1 + (\alpha^{1-p} - 1)\lambda^{-2} z^2)^{-\frac{1}{p-1}} + o(1).$$

Let $k \to +\infty$; the conclusion of (ii) then follows.

(iii) From (4-25) and the fact that $y^+_k/\sqrt{s_k} \to \zeta$ as $k \to +\infty$, we have

$$(\delta^{(k)})^{-2} (y^+_k)^2 = \frac{(\alpha^{1-p} - 1) M^{1-p}}{c_p \lambda^2 (p-1)} \log s_k + o(1).$$

Using (4-24), we then derive

$$(\delta^{(k)})^{-2} (y^+_k)^2 = \frac{2k|\log\lambda|(\alpha^{1-p} - 1) M^{1-p}}{c_p \lambda^2 (p-1)} - \frac{(\alpha^{1-p} - 1) M^{1-p}}{c_p \lambda^2 (p-1)} \log\left(\frac{M^{1-p}\delta^2}{p-1}\right) + o(1),$$

which yields the conclusion of (iii) and finishes the proof of Theorem 4.1. □

δ	0.040	0.020	0.010	0.005
M	0.320	0.160	0.080	0.040

Table 1. The value of M corresponds to the initial data and the initial space step.

4B. ***The numerical results.*** This subsection gives various numerical confirmations for the assertions stated in the previous subsection (Theorem 4.1). All the experiments reported here used $\varphi(x) = 2(1+\cos(\pi x))$ as the initial data, $\alpha = 0.6$ as the parameter for controlling the interval to be refined, $\lambda = \frac{1}{2}$ as the refining factor, $C_\Delta = \frac{1}{4}$ as the stability condition for the scheme (4-3), $p = 3$ and $a = 0.1$, 1 and 10 in the nonlinearity F given in (4-2). The threshold M is chosen to satisfy the condition (4-4). In Table 1, we give some values of M corresponding to the initial data and the initial space step δ. We generally stop the computation after 40 refining phases. Indeed, since $(\delta^{(k)})^{2/(p-1)}\|u^{(k)}(\cdot\,, \tau_k^*)\|_\infty = M$ and $\delta^{(k)} = \lambda\delta^{(k-1)}$, we have by induction that

$$\|u^{(k)}(\cdot\,, \tau_k^*)\|_\infty = (\delta^{(k)})^{-\frac{2}{p-1}} M = (\lambda\delta^{(k-1)})^{-\frac{2}{p-1}} M = \cdots = (\lambda^k\delta)^{-\frac{2}{p-1}} M.$$

With these parameters, we see that the corresponding amplitude of u approaches 10^{12} after 40 iterations.

4B(i). *The value $\tau_k^*/\tau^{(k)}$ tends to a constant as $k \to +\infty$.* It is convenient to denote the computed value of $\tau_k^*/\tau^{(k)}$ by $N^{(k)}$ and the predicted value given in the statement Theorem 4.1(i) by N^{pre}. Note that the value of N^{pre} does not depend on a, but depends on δ because of the relation (4-4). More precisely,

$$N^{\mathrm{pre}}(\delta) = \frac{(1-\lambda^2)\|\varphi\|_\infty^{1-p}}{C_\Delta(p-1)\delta^2}.$$

Then, considering the quantity $N^{(k)}/N^{\mathrm{pre}}$, theoretically it is expected to converge to 1 as k tends to infinity. Table 2 provides computed values of $N^{(k)}/N^{\mathrm{pre}}$ at some selected indices of k, computing with $\delta = 0.005$ and three different values of a. According to the numerical results given in Table 2, the computed values in the cases $a = 10$ and $a = 1.0$ approach to 1 as expected, which gives us a numerical answer for the statement (4-18). However, the numerical results in the case $a = 0.1$ are not good due to the fact that the speed of convergence to the blow-up limit (4-18) is $1/|\log(T-t)|^{a'}$ with $a' = \min\{a, 1\}$ (see Theorem 1.4).

4B(ii). *The function $v^{(k)}(z)$ introduced in Theorem 4.1(ii) converges to a predicted profile as $k \to +\infty$.* As stated in Theorem 4.1(ii), if we plot $v^{(k)}(z)$ over the fixed interval $(-1, 1)$ then the graph of $v^{(k)}$

k	10	15	20	25	30	35	40
$a = 10$	1.0325	1.0203	1.0149	1.0117	1.0096	1.0080	1.0072
$a = 1.0$	0.9699	0.9771	0.9816	0.9845	0.9867	0.9885	0.9899
$a = 0.1$	0.5853	0.5885	0.5923	0.5957	0.5989	0.6016	0.6043

Table 2. The values of $N^{(k)}/N^{\mathrm{pre}}$ at some selected indices of k, computing with $\delta = 0.005$ and three different values of a.

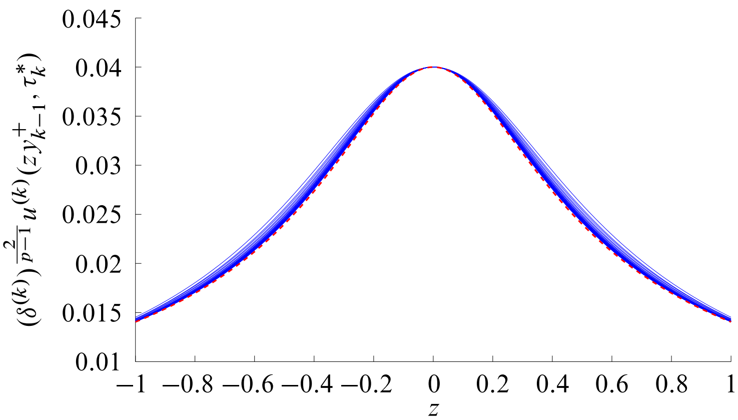

Figure 1. The graph of $v^{(k)}(z)$ at some selected indices of k, computing with $\delta = 0.005$ and $a = 10$. They converge to the predicted profile (the dash line) as stated in (4-15) as k increases.

would converge to the predicted one. Figure 1 gives us a numerical confirmation for this fact, computing with $\delta = 0.005$ and $a = 10$. Looking at Figure 1, we see that the graph of $v^{(k)}$ evidently converges to the predicted one given in the right-hand side of (4-15) as k increases. The last curve $v^{(40)}$ seemingly coincides with the prediction. Figure 2 shows the graph of $v^{(40)}$ and the predicted profile for another experiment with $\delta = 0.005$ and $a = 0.1$. They coincide to within plotting resolution.

In Table 3, we give the error in L^∞ between $v^{(k)}(z)$ at index $k = 40$ and the predicted profile given in the right-hand side of (4-15), namely

$$e_{\delta,a} = \sup_{z\in(-1,1)} \left| v^{(40)}(z) - M(1 + (\alpha^{1-p} - 1)\lambda^{-2} z^2)^{-\frac{1}{p-1}} \right|. \tag{4-26}$$

These numerical computations give us confirmation that the computed profiles v_k converges to the predicted one. Since the error $e_{\delta,a}$ tends to 0 as δ goes to 0, the numerical computations also answer to the stability of the blow-up profile stated in Theorem 1.7(i). In fact, the stability makes the solution visible in numerical simulations.

4B(iii). *The quantity* $(\delta^{(k)})^{-2}(y_k^+)^2$ *behaves like a linear function in* k. For making a quantitative comparison between our numerical results and the predicted behavior as stated in Theorem 4.1(iii), we plot the graph of $(\delta^{(k)})^{-2}(y_k^+)^2$ against k and denote by $\gamma_{\delta,a}$ the slope of this curve. Then, considering

δ	0.04	0.02	0.01	0.005
$a = 10$	0.002906	0.000789	0.000470	0.000238
$a = 1.0$	0.001769	0.000671	0.000359	0.000213
$a = 0.1$	0.002562	0.000687	0.000380	0.000235

Table 3. Error $e_{\delta,a}$ in L^∞ between the computed and predicted profiles, defined in (4-26).

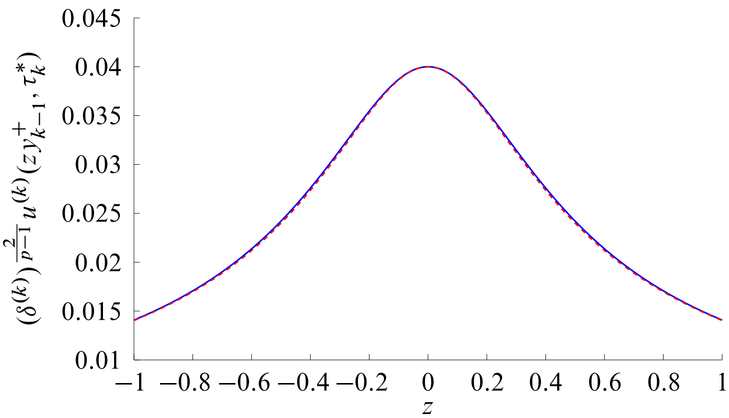

Figure 2. The graph of $v^{(k)}(z)$ at $k = 40$ and the predicted profile given in (4-15), computing with $\delta = 0.005$ and $a = 0.1$. They coincide to within plotting resolution.

the ratio $\gamma_{\delta,a}/\gamma$, where γ is as given in Theorem 4.1(iii). As expected, this ratio $\gamma_{\delta,a}/\gamma$ would approach 1. Figure 3 shows $(\delta^{(k)})^{-2}(y_k^+)^2$ as a function of k, computing with the initial space step $\delta = 0.005$ for different values of a. Looking at Figure 3, we see that the two middle curves, corresponding to the cases $a = 10$ and $a = 1$, behave like the predicted linear function (the top line), while this is not true in the case $a = 0.1$ (the bottom curve). In order to make this clearer, Table 4 lists the values of $\gamma_{\delta,a}/\gamma$, computing with various values of the initial space step δ for three different values of a. Here, the value of $\gamma_{\delta,a}$ is calculated for $20 \le k \le 40$. As Table 4 shows, the numerical values in the cases $a = 10$ and $a = 1$ agree with the prediction stated in Theorem 4.1(ii), while the numerical values in the case $a = 0.1$ are far from the predicted ones.

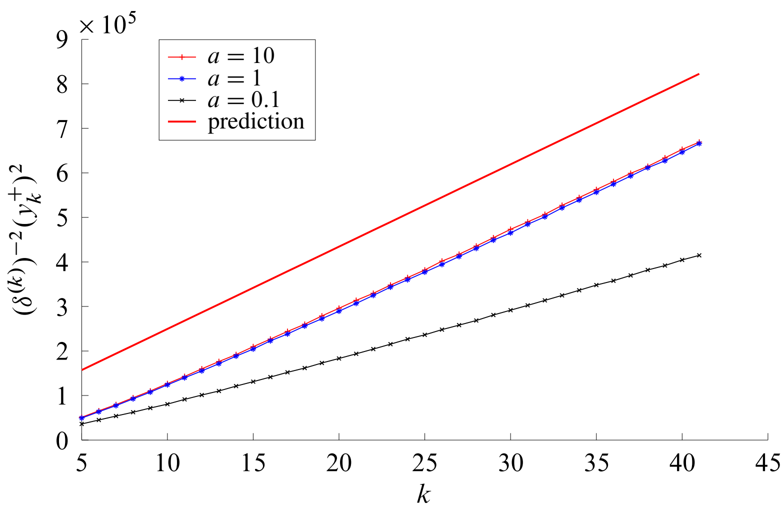

Figure 3. The graph of $(\delta^{(k)})^{-2}(y_k^+)^2$ against k, computing with $\delta = 0.005$ for three different values of a.

δ	0.04	0.02	0.01	0.005
$a=10$	1.9514	1.1541	0.9991	0.9669
$a=1.0$	1.9863	1.1436	1.0052	0.9682
$a=0.1$	1.9538	0.8108	0.6417	0.5986

Table 4. The values of $\gamma_{\delta,a}/\gamma$, computing with various values of the initial space step δ for three different values of a.

Appendix A

The following lemma from [Nguyen 2015] gives the expansion of $\phi(s)$, the unique solution of (1-21) satisfying (1-22):

Lemma A.1. *Let ϕ be a positive solution of the ordinary differential equation*

$$\phi_s = -\frac{\phi}{p-1} + \phi^p + \frac{\mu\phi^p}{\log^a(2+e^{\frac{2s}{p-1}}\phi^2)}.$$

If we assume in addition $\phi(s)\to\kappa$ as $s\to+\infty$, then $\phi(s)$ takes the form

$$\phi(s) = \kappa(1+\eta_a(s))^{-\frac{1}{p-1}} \quad \text{as } s\to+\infty,$$

where

$$\eta_a(s) \sim C_* \int_s^{+\infty} \frac{e^{s-\tau}}{\tau^a}\,d\tau = \frac{C_*}{s^a}\left(1+\sum_{j\geq 1}\frac{b_j}{s^j}\right)$$

with $C_* = \mu\big(\frac{1}{2}(p-1)\big)^a$ and $b_j = (-1)^j\prod_{i=0}^{j-1}(a+i)$.

Proof. See Lemma A.3 in [Nguyen 2015]. □

Appendix B

We aim at proving the following:

Lemma B.1 (estimate of $\omega(s)$). *We have*

$$|\omega(s)| = \mathcal{O}\left(\frac{1}{s^{a+1}}\right) \quad \text{as } s\to+\infty.$$

Proof. From Lemma A.1, we write

$$p(\phi(s)^{p-1}-\kappa^{p-1}) = -\frac{p\eta_a(s)}{p-1}(1+\eta_a(s))^{-1} = -\frac{pC_*}{(p-1)s^a}(1+\eta_a(s))^{-1} + \mathcal{O}\left(\frac{1}{s^{a+1}}\right).$$

A direct calculation yields

$$\begin{aligned} e^{-s}h'(e^{\frac{p}{p-1}}\phi(s)) &= \frac{\mu p\phi^{p-1}(s)}{\log^a(2+e^{\frac{2s}{p-1}}\phi^2(s))} - \frac{2a\mu e^{\frac{2s}{p-1}}\phi^{p+1}(s)}{(2+e^{\frac{2s}{p-1}}\phi^2(s))\log^{a+1}(2+e^{\frac{2s}{p-1}}\phi^2(s))} \\ &= \frac{pC_*}{(p-1)s^a}(1+\eta_a(s))^{-1} + \mathcal{O}\left(\frac{1}{s^{a+1}}\right). \end{aligned}$$

Adding the two above estimates, we obtain the desired result. This ends the proof of Lemma B.1. □

Lemma B.2 (estimate of $R(y,s)$). *We have*

$$|R(y,s)| = \mathcal{O}\left(\frac{|y|^2+1}{s^{a'+1}}\right) \quad \text{as } s \to +\infty$$

with $a' = \min\{1, a\}$.

Proof. Let us write $\varphi(y,s) = (\phi(s)/\kappa)\nu(y,s)$, where

$$\nu(y,s) = \kappa\left(1 + \frac{p-1}{4ps}\sum_{j=1}^{l} y_j^2\right)^{-\frac{1}{p-1}} + \frac{\kappa l}{2ps}.$$

Then, we write $R(y,s) = (\phi(s)/\kappa)R_1(y,s) + R_2(y,s)$, where

$$R_1(y,s) = \nu_s - \Delta\nu - \frac{y}{2}\cdot\nabla\nu - \frac{\nu}{p-1} + \nu^p,$$
$$R_2(y,s) = -\frac{\phi'}{\kappa}\nu - \frac{\phi}{\kappa}\nu^p + \phi^p\left(\frac{\nu}{\kappa}\right)^p + e^{-\frac{ps}{p-1}}h'\left(e^{\frac{s}{p-1}}\frac{\phi\nu}{\kappa}\right).$$

The term $R_1(y,s)$ is already treated in [Velázquez 1992] and it is bounded by

$$|R_1(y,s)| \le \frac{C(|y|^2+1)}{s^2} + C\mathbf{1}_{\{|y|\ge 2K_0\sqrt{s}\}}.$$

To bound R_2, we use the fact that ϕ satisfies (1-22) to write

$$R_2(y,s) = \frac{\nu\phi}{\kappa^p}(\kappa^{p-1} - \phi^{p-1})(\kappa^{p-1} - \nu^{p-1}) + e^{-\frac{ps}{p-1}}\left[h\left(e^{\frac{s}{p-1}}\frac{\phi\nu}{\kappa}\right) - h(e^{\frac{s}{p-1}}\phi)\right] + \left(1 - \frac{\nu}{\kappa}\right)e^{-\frac{ps}{p-1}}h(e^{\frac{s}{p-1}}\phi).$$

Noting that $\nu(y,s) = \kappa + \bar{\nu}(y,s)$ with $|\bar{\nu}(y,s)| \le (C/s)(|y|^2+1)$, uniformly for $y \in \mathbb{R}$ and $s \ge 1$, and recalling from Lemma A.1 that $\phi(s) = \kappa(1+\eta_a(s))^{-1/(p-1)}$, where $\eta_a(s) = \mathcal{O}(s^{-a})$, then using a Taylor expansion, we derive

$$|R_2(y,s)| \le C\left(\frac{|y|^2+1}{s^{a+1}} + \mathbf{1}_{\{|y|\ge 2K_0\sqrt{s}\}}\right).$$

This concludes the proof of Lemma B.2. □

Acknowledgements

The authors are grateful to M. A. Hamza for several helpful conversations pertaining to this work. They would also like to thank the referees for their valuable comments, which helped to improve the paper.

References

[Abia, López-Marcos and Martínez 1998] L. M. Abia, J. C. López-Marcos, and J. Martínez, "On the blow-up time convergence of semidiscretizations of reaction–diffusion equations", *Appl. Numer. Math.* **26**:4 (1998), 399–414. MR 99b:65108 Zbl 0929.65070

[Abia, López-Marcos and Martínez 2001] L. M. Abia, J. C. López-Marcos, and J. Martínez, "The Euler method in the numerical integration of reaction–diffusion problems with blow-up", *Appl. Numer. Math.* **38**:3 (2001), 287–313. MR 2002k:65114 Zbl 0988.65076

[Ball 1977] J. M. Ball, "Remarks on blow-up and nonexistence theorems for nonlinear evolution equations", *Quart. J. Math. Oxford Ser.* (2) **28**:112 (1977), 473–486. MR 57 #13150 Zbl 0377.35037

[Baruch et al. 2010] G. Baruch, G. Fibich, and N. Gavish, "Singular standing-ring solutions of nonlinear partial differential equations", *Phys. D* **239**:20-22 (2010), 1968–1983. MR 2011h:35267 Zbl 1203.35259

[Bebernes and Bricher 1992] J. Bebernes and S. Bricher, "Final time blowup profiles for semilinear parabolic equations via center manifold theory", *SIAM J. Math. Anal.* **23**:4 (1992), 852–869. MR 93h:35091 Zbl 0754.35055

[Berger and Kohn 1988] M. Berger and R. V. Kohn, "A rescaling algorithm for the numerical calculation of blowing-up solutions", *Comm. Pure Appl. Math.* **41**:6 (1988), 841–863. MR 89g:65154 Zbl 0652.65070

[Bressan 1990] A. Bressan, "On the asymptotic shape of blow-up", *Indiana Univ. Math. J.* **39**:4 (1990), 947–960. MR 91i:35030 Zbl 0705.35014

[Bressan 1992] A. Bressan, "Stable blow-up patterns", *J. Differential Equ.* **98**:1 (1992), 57–75. MR 93c:35041 Zbl 0770.35010

[Bricmont and Kupiainen 1994] J. Bricmont and A. Kupiainen, "Universality in blow-up for nonlinear heat equations", *Nonlinearity* **7**:2 (1994), 539–575. MR 95h:35030 Zbl 0857.35018

[Cangiani et al. $\geq$ 2016] A. Cangiani, E. H. Georgoulis, I. Kyza, and S. Metcalfe, "Adaptivity and blow-up detection for nonlinear non-stationary convection-diffusion problems", in preparation.

[Fermanian Kammerer and Zaag 2000] C. Fermanian Kammerer and H. Zaag, "Boundedness up to blow-up of the difference between two solutions to a semilinear heat equation", *Nonlinearity* **13**:4 (2000), 1189–1216. MR 2001g:35123 Zbl 0954.35085

[Fermanian Kammerer et al. 2000] C. Fermanian Kammerer, F. Merle, and H. Zaag, "Stability of the blow-up profile of non-linear heat equations from the dynamical system point of view", *Math. Ann.* **317**:2 (2000), 347–387. MR 2001d:35088 Zbl 0971.35038

[Filippas and Kohn 1992] S. Filippas and R. V. Kohn, "Refined asymptotics for the blowup of $u_t - \Delta u = u^p$", *Comm. Pure Appl. Math.* **45**:7 (1992), 821–869. MR 93g:35066 Zbl 0784.35010

[Filippas and Liu 1993] S. Filippas and W. X. Liu, "On the blowup of multidimensional semilinear heat equations", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **10**:3 (1993), 313–344. MR 94k:35139 Zbl 0815.35039

[Fujita 1966] H. Fujita, "On the blowing up of solutions of the Cauchy problem for $u_t = \Delta u + u^{1+\alpha}$", *J. Fac. Sci. Univ. Tokyo Sect. I* **13** (1966), 109–124. MR 35 #5761 Zbl 0163.34002

[Giga and Kohn 1987] Y. Giga and R. V. Kohn, "Characterizing blowup using similarity variables", *Indiana Univ. Math. J.* **36**:1 (1987), 1–40. MR 88c:35021 Zbl 0601.35052

[Giga and Kohn 1989] Y. Giga and R. V. Kohn, "Nondegeneracy of blowup for semilinear heat equations", *Comm. Pure Appl. Math.* **42**:6 (1989), 845–884. MR 90k:35034

[Giga et al. 2004a] Y. Giga, S. Matsui, and S. Sasayama, "Blow up rate for semilinear heat equations with subcritical nonlinearity", *Indiana Univ. Math. J.* **53**:2 (2004), 483–514. MR 2005g:35153 Zbl 1058.35096

[Giga et al. 2004b] Y. Giga, S. Matsui, and S. Sasayama, "On blow-up rate for sign-changing solutions in a convex domain", *Math. Methods Appl. Sci.* **27**:15 (2004), 1771–1782. MR 2005g:35154 Zbl 1066.35043

[Groisman 2006] P. Groisman, "Totally discrete explicit and semi-implicit Euler methods for a blow-up problem in several space dimensions", *Computing* **76**:3–4 (2006), 325–352. MR 2006m:65181 Zbl 1087.65093

[Groisman and Rossi 2001] P. Groisman and J. D. Rossi, "Asymptotic behaviour for a numerical approximation of a parabolic problem with blowing up solutions", *J. Comput. Appl. Math.* **135**:1 (2001), 135–155. MR 2002g:35101 Zbl 0991.65090

[Groisman and Rossi 2004] P. Groisman and J. D. Rossi, "Dependence of the blow-up time with respect to parameters and numerical approximations for a parabolic problem", *Asymptot. Anal.* **37**:1 (2004), 79–91. MR 2005a:35131 Zbl 1047.35064

[Hamza and Zaag 2012a] M. A. Hamza and H. Zaag, "Lyapunov functional and blow-up results for a class of perturbations of semilinear wave equations in the critical case", *J. Hyperbolic Differ. Equ.* **9**:2 (2012), 195–221. MR 2928106 Zbl 1255.35171

[Hamza and Zaag 2012b] M. A. Hamza and H. Zaag, "A Lyapunov functional and blow-up results for a class of perturbed semilinear wave equations", *Nonlinearity* **25**:9 (2012), 2759–2773. MR 2967123 Zbl 1255.35057

[Herrero and Velázquez 1992a] M. A. Herrero and J. J. L. Velázquez, "Blow-up profiles in one-dimensional, semilinear parabolic problems", *Comm. Partial Differential Equations* **17**:1–2 (1992), 205–219. MR 93b:35066 Zbl 0772.35027

[Herrero and Velázquez 1992b] M. A. Herrero and J. J. L. Velázquez, "Comportement générique au voisinage d'un point d'explosion pour des solutions d'équations paraboliques unidimensionnelles", *C. R. Acad. Sci. Paris Sér. I Math.* **314**:3 (1992), 201–203. MR 92m:35140 Zbl 0765.35009

[Herrero and Velázquez 1992c] M. A. Herrero and J. J. L. Velázquez, "Flat blow-up in one-dimensional semilinear heat equations", *Differential Integral Equations* **5**:5 (1992), 973–997. MR 93d:35065 Zbl 0767.35036

[Herrero and Velázquez 1992d] M. A. Herrero and J. J. L. Velázquez, "Generic behaviour of one-dimensional blow up patterns", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **19**:3 (1992), 381–450. MR 94b:35048 Zbl 0798.35081

[Herrero and Velázquez 1993] M. A. Herrero and J. J. L. Velázquez, "Blow-up behaviour of one-dimensional semilinear parabolic equations", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **10**:2 (1993), 131–189. MR 94g:35030 Zbl 0813.35007

[Kyza and Makridakis 2011] I. Kyza and C. Makridakis, "Analysis for time discrete approximations of blow-up solutions of semilinear parabolic equations", *SIAM J. Numer. Anal.* **49**:1 (2011), 405–426. MR 2012d:65185 Zbl 1227.65081

[Levine 1973] H. A. Levine, "Some nonexistence and instability theorems for solutions of formally parabolic equations of the form $Pu_t = -Au + \mathscr{F}(u)$", *Arch. Rational Mech. Anal.* **51** (1973), 371–386. MR 50 #714 Zbl 0278.35052

[Masmoudi and Zaag 2008] N. Masmoudi and H. Zaag, "Blow-up profile for the complex Ginzburg–Landau equation", *J. Funct. Anal.* **255**:7 (2008), 1613–1666. MR 2010a:35246 Zbl 1158.35016

[Merle and Zaag 1997] F. Merle and H. Zaag, "Stability of the blow-up profile for equations of the type $u_t = \Delta u + |u|^{p-1}u$", *Duke Math. J.* **86**:1 (1997), 143–195. MR 98d:35098 Zbl 0872.35049

[N'gohisse and Boni 2011] F. K. N'gohisse and T. K. Boni, "Numerical blow-up for a nonlinear heat equation", *Acta Math. Sin. (Engl. Ser.)* **27**:5 (2011), 845–862. MR 2012d:35192 Zbl 1221.35075

[Nguyen 2014] V. T. Nguyen, "Numerical analysis of the rescaling method for parabolic problems with blow-up in finite time", preprint, 2014. arXiv 1403.7547

[Nguyen 2015] V. T. Nguyen, "On the blow-up results for a class of strongly perturbed semilinear heat equations", *Discrete Contin. Dyn. Syst.* **35**:8 (2015), 3585–3626. MR 3320139

[Nguyen and Zaag 2014] V. T. Nguyen and H. Zaag, "Construction of a stable blow-up solution for a class of strongly perturbed semilinear heat equations", preprint, 2014. To appear in *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* arXiv 1406.5233

[Velázquez 1992] J. J. L. Velázquez, "Higher-dimensional blow up for semilinear parabolic equations", *Comm. Partial Differential Equations* **17**:9–10 (1992), 1567–1596. MR 93k:35044 Zbl 0813.35009

[Velázquez 1993] J. J. L. Velázquez, "Classification of singularities for blowing up solutions in higher dimensions", *Trans. Amer. Math. Soc.* **338**:1 (1993), 441–464. MR 93j:35101 Zbl 0803.35015

[Weissler 1981] F. B. Weissler, "Existence and nonexistence of global solutions for a semilinear heat equation", *Israel J. Math.* **38**:1–2 (1981), 29–40. MR 82g:35059 Zbl 0476.35043

Received 25 Nov 2014. Revised 20 Aug 2015. Accepted 11 Oct 2015.

VAN TIEN NGUYEN: `tien.nguyen@nyu.edu`
Department of Mathematics, New York University Abu Dhabi, Saadiyat Island, PO Box 129188, Abu Dhabi, United Arab Emirates

HATEM ZAAG: `Hatem.Zaag@univ-paris13.fr`
LAGA, CNRS (UMR 7539), Université Paris 13, Sorbonne Paris Cité, 93430 Villetaneuse, France

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. 1 2016

Rademacher functions in Nakano spaces — 1
SERGEY ASTASHKIN and MIECZYSŁAW MASTYŁO

Nonexistence of small doubly periodic solutions for dispersive equations — 15
DAVID M. AMBROSE and J. DOUGLAS WRIGHT

The borderlines of invisibility and visibility in Calderón's inverse problem — 43
KARI ASTALA, MATTI LASSAS and LASSI PÄIVÄRINTA

A characterization of 1-rectifiable doubling measures with connected supports — 99
JONAS AZZAM and MIHALIS MOURGOGLOU

Construction of Hadamard states by characteristic Cauchy problem — 111
CHRISTIAN GÉRARD and MICHAŁ WROCHNA

Global-in-time Strichartz estimates on nontrapping, asymptotically conic manifolds — 151
ANDREW HASSELL and JUNYONG ZHANG

Limiting distribution of elliptic homogenization error with periodic diffusion and random potential — 193
WENJIA JING

Blow-up results for a strongly perturbed semilinear heat equation: theoretical analysis and numerical method — 229
VAN TIEN NGUYEN and HATEM ZAAG

2157-5045(2016)9:1;1-B