Analysis & PDE

msp

Analysis & PDE

msp.org/apde

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

FLAG HARDY SPACES AND MARCINKIEWICZ MULTIPLIERS ON THE HEISENBERG GROUP

YONGSHENG HAN, GUOZHEN LU AND ERIC SAWYER

Marcinkiewicz multipliers are L^p bounded for $1 < p < \infty$ on the Heisenberg group $\mathbb{H}^n \simeq \mathbb{C}^n \times \mathbb{R}$, as shown by D. Müller, F. Ricci, and E. M. Stein. This is surprising in that these multipliers are invariant under a two-parameter group of dilations on $\mathbb{C}^n \times \mathbb{R}$, while there is *no* two-parameter group of *automorphic* dilations on $\mathbb{H}^n$. This lack of automorphic dilations underlies the failure of such multipliers to be in general bounded on the classical Hardy space H^1 on the Heisenberg group, and also precludes a pure product Hardy space theory.

We address this deficiency by developing a theory of *flag* Hardy spaces H^p_{flag} on the Heisenberg group, $0 < p \le 1$, that is in a sense "intermediate" between the classical Hardy spaces H^p and the product Hardy spaces H^p_{product} on $\mathbb{C}^n \times \mathbb{R}$ developed by A. Chang and R. Fefferman. We show that flag singular integral operators, which include the aforementioned Marcinkiewicz multipliers, are bounded on H^p_{flag}, as well as from H^p_{flag} to L^p, for $0 < p \le 1$. We also characterize the dual spaces of H^1_{flag} and H^p_{flag}, and establish a Calderón–Zygmund decomposition that yields standard interpolation theorems for the flag Hardy spaces H^p_{flag}. In particular, this recovers some L^p results of Müller, Ricci, and Stein (but not their sharp versions) by interpolating between those for H^p_{flag} and L^2.

1. Introduction

Classical Calderón–Zygmund theory centers around singular integrals associated with the Hardy–Little-wood maximal operator M that commutes with the usual dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta x_1, \ldots, \delta x_n)$ for $\delta > 0$. On the other hand, *product* Calderón–Zygmund theory centers around singular integrals associated with the *strong* maximal function M_S that commutes with the multiparameter dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta_1 x_1, \ldots, \delta_n x_n)$ for $\delta = (\delta_1, \ldots, \delta_n) \in \mathbb{R}^n_+$. The strong maximal function [Jessen et al. 1935] is given by

$$M_S(f)(x) = \sup_{x \in R} \frac{1}{|R|} \int_R |f(y)| \, dy, \tag{1-1}$$

where the supremum is taken over the family of all rectangles R with sides parallel to the axes.

For Calderón–Zygmund theory in the product setting, one considers operators of the form $Tf = K * f$, where K is homogeneous, that is, $\delta_1 \cdots \delta_n K(\delta \cdot x) = K(x)$, or, more generally, $K(x)$ satisfies certain differential inequalities and cancellation conditions such that the kernels $\delta_1 \cdots \delta_n K(\delta \cdot x)$ also satisfy the same bounds. Such operators have been studied, for example, in [Gundy and Stein 1979; Fefferman and Stein 1982; Fefferman 1986; 1987; 1999; Chang 1979; Chang and Fefferman 1985; 1982; 1980; Journé 1985; 1986; Pipher 1986; Ferguson and Lacey 2002], where both the L^p theory for $1 < p < \infty$ and H^p theory for $0 < p \leq 1$ were developed. More precisely, Fefferman and Stein [1982] studied the L^p boundedness ($1 < p < \infty$) for the product convolution singular integral operators. Journé [1985; 1988] introduced non-convolution-product singular integral operators, established the product $T1$ theorem, and proved the $L^\infty \to$ BMO boundedness of such operators. The product Hardy space $H^p(\mathbb{R}^n \times \mathbb{R}^m)$ was first introduced by Gundy and Stein [1979]. Chang and Fefferman [1985; 1982; 1980] developed the theory of atomic decomposition and established the dual space of the Hardy space $H^1(\mathbb{R}^n \times \mathbb{R}^m)$, namely the product BMO($\mathbb{R}^n \times \mathbb{R}^m$) space. Another characterization of such product BMO space was given in conjunction with Hankel theorems and commutators in the product setting by Ferguson and Lacey [2002] and Lacey and Terwilleger [2005]. Carleson [1974] disproved by a counterexample the conjecture that the product atomic Hardy space on $\mathbb{R}^n \times \mathbb{R}^m$ could be defined by rectangle atoms. This motivated Chang and Fefferman to replace the role of cubes in the classical atomic decomposition of $H^p(\mathbb{R}^n)$ by arbitrary open sets of finite measures in the product $H^p(\mathbb{R}^n \times \mathbb{R}^m)$. Subsequently, Fefferman [1987] established the criterion for the $H^p \to L^p$ boundedness of singular integral operators in Journé's class by considering its action only on rectangle atoms by using Journé's lemma. However, Fefferman's criterion cannot be extended to three or more parameters without further assumptions on the nature of T, as shown in [Journé 1985; Journé 1988]. In fact, Journé provided a counterexample in the three-parameter setting of singular integral operators such that Fefferman's criterion breaks down. Subsequently, the H^p to L^p boundedness for Journé's class of singular integral operators with arbitrary number of parameters was established by J. Pipher [1986] by considering directly the action of the operator on (nonrectangle) atoms and an extension of Journé's geometric lemma to higher dimensions.

On the other hand, multiparameter analysis has only recently been developed for L^p theory with $1 < p < \infty$ when the underlying multiparameter structure is not explicit, but *implicit*, as in the flag multiparameter structure studied in [Nagel et al. 2001] and its counterpart on the Heisenberg group $\mathbb{H}^n$

studied in [Müller et al. 1995; 1996]. In these latter two papers the authors obtained the surprising result that certain Marcinkiewicz multipliers, invariant under a two-parameter group of dilations on $\mathbb{C}^n \times \mathbb{R}$, are bounded on $L^p(\mathbb{H}^n)$, *despite* the absence of a two-parameter automorphic group of dilations on $\mathbb{H}^n$. This striking result exploited an implicit product, or *semiproduct,* structure underlying the group multiplication in $\mathbb{H}^n \simeq \mathbb{C}^n \times \mathbb{R}$. In contrast to this, it is not hard to see that the class of flag singular integrals considered there is *not* in general bounded on the standard one-parameter Hardy space $H^1(\mathbb{H}^n)$ as in [Fefferman and Stein 1972] (see, for example, Theorem 67 in Section 11 below). The lesson learned here is that Hardy space theories for $0 < p \le 1$ must be tailored to the invariance properties of the class of singular integral operators under consideration.

The goal of this paper is to develop for the Heisenberg group a theory of *flag* Hardy spaces H_{flag}^p with $0 < p \le 1$. The first two authors have treated the Euclidean flag structure in [Han and Lu 2008]; see also the multiparameter setting associated with the Zygmund dilation [Han and Lu 2010]. The ideas developed in this paper and [Han and Lu 2008; Han and Lu 2010] have been adapted to some other multiparameter cases, such as the product spaces of Carnot–Carathéodory spaces [Han et al. 2013a], where the L^p theory was established in [Nagel and Stein 2004], and the composition of two singular integrals with different homogeneity [Han et al. 2013b].

This flag theory for the Heisenberg group is most conveniently explained when $p = 1$ in the more general context of spaces $(X, \rho, d\mu)$ of homogeneous type [Coifman and Weiss 1976], which already include Euclidean spaces $\mathbb{R}^N$ and stratified graded nilpotent Lie groups such as the Heisenberg groups $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$. We may assume here that ρ and $d\mu$ are connected by the equivalence

$$\mu(B_\rho(x, r)) \approx r, \quad \text{where } B_\rho(x, r) = \{y \in X : \rho(x, y) < r\}. \tag{1-2}$$

In particular, the usual structure on Euclidean space $\mathbb{R}^n$ is given by $\rho(x, y) = |x - y|^n$ and $d\mu(x) = dx$.

Recall that one of several equivalent definitions of the Hardy space $H^1(X)$ is as the set of $f \in (C^\eta(X))^*$ with

$$\|f\|_{H^1(X)} \equiv \|g(f)\|_{L^1(d\mu)} < \infty,$$

where the Littlewood–Paley g-function $g(f)$ is given by

$$g(f) = \left\{ \sum_{j=-\infty}^{\infty} |E_j f|^2 \right\}^{\frac{1}{2}},$$

where $\{E_j\}_{j=-\infty}^{\infty}$ is an appropriate Littlewood–Paley decomposition of the identity on $L^2(d\mu)$.

The *product* Hardy space $H_{\mathrm{product}}^1(X \times X')$ corresponding to a product of homogeneous spaces $(X, \rho, d\mu)$ and $(X', \rho', d\mu')$ is given as the set of $f \in (C^\eta(X \times X'))^*$ with

$$\|f\|_{H_{\mathrm{product}}^1(X \times X')} \equiv \|g_{\mathrm{product}}(f)\|_{L^1(d\mu \times d\mu')} < \infty,$$

where the *product* Littlewood–Paley g-function $g_{\mathrm{product}}(f)$ is given by

$$g_{\mathrm{product}}(f) = \left\{ \sum_{j,j'=-\infty}^{\infty} |D_j D'_{j'} f|^2 \right\}^{\frac{1}{2}},$$

and where $\{D_j\}_{j=-\infty}^{\infty}$ and $\{D'_{j'}\}_{j'=-\infty}^{\infty}$ are Littlewood–Paley decompositions of the identities on $L^2(d\mu)$ and $L^2(d\mu')$, respectively (and act separately on the respective distinct variables). Note that if $j = j'$, then $D_j D'_{j'} = D_j D'_j$ satisfies estimates similar to those for E_j in the standard *one-parameter* Hardy space $H^1(X \times X')$. Thus, we see that

$$g_{\text{product}}(f) = \left\{ \sum_{j,j'=-\infty}^{\infty} |D_j D'_{j'} f|^2 \right\}^{\frac{1}{2}} \geq \left\{ \sum_j |D_j D'_j f|^2 \right\}^{\frac{1}{2}} \approx \left\{ \sum_j |E_j f|^2 \right\}^{\frac{1}{2}} = g(f),$$

and so we have the inclusion

$$H^1_{\text{product}}(X \times X') \subset H^1(X \times X').$$

Now we specialize the space of homogeneous type X to be the Heisenberg group $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$. The *flag* structure on the Heisenberg group $\mathbb{H}^n$ arises in an intermediate manner, namely, as a homogeneous space structure derived from the Heisenberg multiplication law that is adapted to the product of the homogeneous spaces $\mathbb{C}^m$ and $\mathbb{R}$. The appropriate definition of the flag Hardy space $H^1_{\text{flag}}(\mathbb{H}^n)$ is already suggested in [Müller et al. 1996], where a Littlewood–Paley g-function g_{flag} is introduced that is adapted to the flag structure on the Heisenberg group $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$:

$$g_{\text{flag}}(f) = \left\{ \sum_{j,k=-\infty}^{\infty} |E_k D_j f|^2 \right\}^{\frac{1}{2}},$$

where $\{D_j\}_{j=-\infty}^{\infty}$ is the standard Littlewood–Paley decomposition of the identity on $L^2(\mathbb{H}^n)$, and $\{E_k\}_{k=-\infty}^{\infty}$ is the standard Littlewood–Paley decomposition of the identity on $L^2(\mathbb{R})$. One can then define $H^1_{\text{flag}}(\mathbb{H}^n)$ to consist of appropriate "distributions" f on $\mathbb{H}^n$ with

$$\|f\|_{H^1_{\text{flag}}(\mathbb{H}^n)} \equiv \|g_{\text{flag}}(f)\|_{L^1(\mathbb{H}^n)} < \infty.$$

Now, for $k \leq 2j$, it turns out that $E_k D_j$ is essentially the *one-parameter* Littlewood–Paley function D_j; while, for $k > 2j$, it turns out that $E_k D_j$ is essentially the *product* Littlewood–Paley function $E_k F_j$, where $\{F_j\}_{j=-\infty}^{\infty}$ is the standard Littlewood–Paley decomposition of the identity on $L^2(\mathbb{C}^n)$. Thus we see that $g_{\text{flag}}(f)$ is a *semiproduct* Littlewood–Paley function satisfying

$$g_{\text{product}}(f) \gtrsim g_{\text{flag}}(f) \gtrsim g(f), \quad H^1_{\text{product}}(X \times X') \subset H^1_{\text{flag}}(X \times X') \subset H^1(X \times X').$$

We describe this structure as "semiproduct", since only *vertical* Heisenberg rectangles (which are essentially unions of contiguous Heisenberg balls of fixed radius stacked one on top of the other) arise essentially as the supports of the components $E_k D_j$, when $k > 2j$. When $k \leq 2j$, the support of $E_k D_j$ is essentially a Heisenberg cube. Thus no *horizontal* rectangles arise, and the structure is "semiproduct".

Of course, we must also address the nature of the "distributions" referred to above, and for this we will use a lifting technique introduced in [Müller et al. 1995] to define *projected* flag molecular spaces $\mathcal{M}_{\text{flag}}(\mathbb{H}^n)$, and then the aforementioned distributions will be elements of the dual space $\mathcal{M}_{\text{flag}}(\mathbb{H}^n)'$. We also show that these distributions are essentially the same as those obtained from the dual of a more familiar *moment* flag molecular space $\mathcal{M}_F(\mathbb{H}^n)$. Finally, we mention that a theory of flag Hardy spaces

can also be developed with the techniques used here, but without recourse to any notion of "distributions", by simply defining $H^p_{\text{abstract}}(\mathbb{H}^n)$ to be the abstract completion of the metric space

$$X^p(\mathbb{H}^n) \equiv \{f \in L^2(\mathbb{H}^n) : g_{\text{flag}}(f) \in L^p(\mathbb{H}^n)\}$$

with metric

$$d(f_1, f_2) \equiv \|g_{\text{flag}}(f_1 - f_2)\|^p_{L^p(\mathbb{H}^n)}, \quad f_j \in X^p(\mathbb{H}^n).$$

We show that the abstract space $H^p_{\text{abstract}}(\mathbb{H}^n)$, whose elements are realized only as equivalence classes of Cauchy sequences, is in fact isomorphic to the space $H^p_{\text{flag}}(\mathbb{H}^n)$, whose elements have the advantage of being realized as a subspace of distributions, namely those f in $\mathcal{M}_{\text{flag}}(\mathbb{H}^n)'$ whose flag Littlewood–Paley function $g_{\text{flag}}(f)$ belongs to $L^p(\mathbb{H}^n)$. Here $\mathcal{M}_{\text{flag}}(\mathbb{H}^n)$ is a molecule space with implicit product structure.

In Part I of the paper we define flag Hardy spaces and state our results. In Part II we give the proofs, and in Part III we construct a dyadic grid adapted to the flag structure.

Remark 1. Some of the proofs we need in this paper are straightforward modifications of arguments already in the literature, and in order not to interrupt the flow of the paper, we have left these proofs out. However, all the details are included in the expanded version of this paper [Han et al. 2012].

Part I. Flag Hardy spaces: definitions and results

Our point of departure is to develop a *wavelet* Calderón reproducing formula associated with the given two-parameter structure as in [Müller et al. 1996], and then to prove a Plancherel–Pólya-type inequality in this setting. This will provide the flexibility needed to define flag Hardy spaces and prove boundedness of flag singular integrals, duality, and interpolation theorems for these spaces. To explain the novelty in this approach more carefully, we point out the following three types of reproducing formulas derived from the original idea of Calderón:

$$f(x) = \int_0^\infty \psi_t * \psi_t * f(x) \frac{dt}{t},$$

$$f(x) = \sum_{j \in \mathbb{Z}} \widetilde{D}_j D_j f(x),$$

$$f(x) = \sum_j \sum_I \{|I|(\psi_j * f)(x_I)\} \tilde{\psi}_j(x, x_I).$$

We refer to the first formula as a *continuous* Calderón reproducing formula, its advantage being the use of *compactly* supported components ψ_t that are repeated. We refer to the second formula as a *discrete* Calderón reproducing formula, in which D_j is generally a compactly supported *nonconvolution* operator in a space of homogeneous type, and $\widetilde{D}_j$ is no longer compactly supported but satisfies *molecular* estimates. In certain cases, such as in Euclidean space, it is possible to use the Fourier transform to obtain a discrete decomposition with repeated convolution operators $D_j = \psi_j$.

Finally, we refer to the third formula as a *wavelet* Calderón reproducing formula, which can also be developed in a space of homogeneous type. For example, such formulas were first developed in certain situations in [Frazier and Jawerth 1990]. The advantage of the third formula is that it expresses f as a

sum of molecular, or wavelet-like, functions $\tilde{\psi}_j(x, x_I)$ with coefficients $|I|(\psi_j * f)(x_I)$ that are obtained by evaluating $\psi_j * f$ at *any* convenient point in the set I from a dyadic decomposition at scale 2^j of the space. As a consequence, we can replace the coefficient $|I|(\psi_j * f)(x_I)$ with either the supremum or infimum of such choices and retain appropriate estimates(see Theorem 19 below). We note in passing that the collection of functions $\{\tilde{\psi}_j(x, x_I)\}_{j,I}$ forms a *Riesz* basis for L^2. In certain cases when such functions form an *orthogonal* basis, the decomposition is referred to as a *wavelet* decomposition, and it is from this that we borrow our terminology.

This "wavelet" scheme is particularly useful in dealing with the Hardy spaces H^p for $0 < p \leq 1$, and using this, we will show that flag singular integral operators are bounded on H_{flag}^p for all $0 < p \leq 1$, and furthermore that these operators are bounded from H_{flag}^p to L^p for all $0 < p \leq 1$. These ideas can also be applied in the pure product setting to provide a different approach to proving H_{product}^p to L^p boundedness than that used by Fefferman, and thus to bypass both the action of singular integral operators on rectangle atoms, and the use of Journé's covering lemma.

We now recall the example of implicit multiparameter structure that provides the main motivation for this paper. In [Müller et al. 1995], Müller, Ricci, and Stein uncovered a new class of flag singular integrals on Heisenberg(-type) groups, which arose in the investigation of Marcinkiewicz multipliers. To be more precise, let $m(\mathcal{L}, iT)$ be the Marcinkiewicz multiplier operator, where $\mathcal{L}$ is the sub-Laplacian, T is the central element of the Lie algebra on the Heisenberg group $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$, and m satisfies Marcinkiewicz conditions as in [Müller et al. 1995]. It was proved in [Müller et al. 1995] that the kernel of $m(\mathcal{L}, iT)$ satisfies the standard one-parameter Calderón–Zygmund-type estimates associated with automorphic dilations in the region where $|t| < |z|^2$, and the multiparameter Calderón–Zygmund-type estimates in the region where $|t| \geq |z|^2$.

The proof of L^p boundedness of $m(\mathcal{L}, iT)$ given in [Müller et al. 1995] requires lifting the operator to a larger group, $\mathbb{H}^n \times \mathbb{R}$. This lifts K, the kernel of $m(\mathcal{L}, iT)$ on $\mathbb{H}^n$, to a *product* kernel $\tilde{K}$ on $\mathbb{H}^n \times \mathbb{R}$. The lifted kernel $\tilde{K}$ is constructed so that it projects to K by

$$K(z, t) = \int_{-\infty}^{\infty} \tilde{K}(z, t - u, u)\, du,$$

taken in the sense of distributions. The operator $\tilde{T}$ corresponding to the product kernel $\tilde{K}$ can be dealt with in terms of tensor products of operators, and one can obtain their L^p boundedness from the known pure product theory. Finally, the L^p boundedness of the operator with kernel K follows from the transference method of [Coifman and Weiss 1976], using the projection $\pi : \mathbb{H}^n \times \mathbb{R} \to \mathbb{H}^n$ given by $\pi((z, t), u) = (z, t + u)$. One of our main results, Corollary 27 below, is an extension of the boundedness of $m(\mathcal{L}, iT)$ to flag Hardy spaces H_{flag}^p for all $0 < p \leq 1$, and follows from the boundedness of flag singular integrals on H_{flag}^p.

In [Müller et al. 1996], the authors obtained the same boundedness results, but with optimal regularity on the multipliers. This required working directly on the group without lifting to a product, and led to the introduction of a *continuous* flag Littlewood–Paley g-function and a corresponding *continuous* Calderón reproducing formula. We remark that one of the main features of our extension of these results to H^p for $0 < p \leq 1$ is the construction of a *wavelet* Calderón reproducing formula.

We note that the regularity satisfied by flag singular kernels is better than that of the product singular kernels. More precisely, the singularity of the standard pure product kernel on $\mathbb{C}^n \times \mathbb{R}$ is contained in the union $\{(z, 0)\} \cup \{(0, u)\}$ of two subspaces, while the singularity of $K(z, u)$, the flag singular kernel on $\mathbb{H}^n \times \mathbb{R}$ defined by Definition 7 below, is contained in a single subspace $\{(0, u)\}$, but is more singular on yet a smaller subspace $\{(0, 0)\}$, a situation described neatly in terms of the flag (or filtration) of subspaces, $\{(0, 0)\} \subsetneq \{(0, u)\} \subsetneq \mathbb{H}^n$. In the following, we describe some natural questions that arise.

Question 1. What is the correct definition of a flag Hardy space H_{flag}^p associated with flag singular integral operators for $0 < p \le 1$ so that both (1) flag singular integral operators are bounded, and (2) a satisfactory theory of interpolation emerges?

Question 2. What is the correct definition of spaces $\mathrm{BMO}_{\mathrm{flag}}$ of bounded mean oscillation for flag singular integral operators, and are the singular integrals bounded on them?

Question 3. What is the duality theory for H_{flag}^p? Is there an analogue of BMO and Carleson measure-type function spaces which are dual spaces of the flag Hardy spaces H_{flag}^p as in the pure product setting?

Question 4. Is there a Calderón–Zygmund decomposition adapted to functions in flag Hardy spaces H_{flag}^p that leads, for example, to an appropriate theory of interpolation?

Question 5. What is the relationship between classical Hardy spaces H^p and the flag Hardy spaces H_{flag}^p?

We address these five questions as follows. As in the L^p theory for $p > 1$ considered in [Müller et al. 1995], one is naturally tempted to establish Hardy space theory under the implicit two-parameter structure associated with the flag singular kernel by invoking the method of lifting to the pure product setting together with the transference method in [Coifman and Weiss 1976]. However, this direct lifting method is not readily adaptable to the case of $p \le 1$ because the transference method is not known to be valid. A different approach centering on the use of a continuous flag Littlewood–Paley g-function was carried out in [Müller et al. 1996]. This suggests that the flag Hardy space H_{flag}^p associated with this implicit two-parameter structure for $0 < p \le 1$ should be defined in terms of this or a similar g-function. Crucial for this is the use of a space of test functions arising from the lifting technique in [Müller et al. 1995], and a "wavelet" Calderón reproducing formula adapted to these test functions. Here is the order in which we implement these ideas.

(1) We first use the L^p theory of Littlewood–Paley square functions g_{flag} as in [Müller et al. 1996] to develop a Plancherel–Pólya-type inequality.

(2) We next define the flag Hardy spaces H_{flag}^p using the flag g-function g_{flag} together with a space of test functions that is motivated by the lifting technique in [Müller et al. 1995]. We then develop the theory of Hardy spaces H_{flag}^p associated to the two-parameter flag structures and the boundedness of flag singular integrals on these spaces. We also establish the boundedness of flag singular integrals from H_{flag}^p to L^p.

(3) We then turn to duality theory for the flag Hardy space H_{flag}^p and introduce the dual space $\mathrm{CMO}_{\mathrm{flag}}^p$. In particular we establish the duality between H_{flag}^1 and the space $\mathrm{BMO}_{\mathrm{flag}}$. We then establish the

boundedness of flag singular integrals on $\mathrm{BMO}_{\mathrm{flag}}$. It is worthwhile to point out that in the classical one-parameter or pure product case, BMO is related to the concept of Carleson measure. The space $\mathrm{CMO}^p_{\mathrm{flag}}$ for all $0 < p \le 1$, as the dual space of H^p_{flag} introduced in this paper, is then defined by a generalized Carleson measure condition.

(4) We finally establish a Calderón–Zygmund decomposition lemma for any H^p_{flag} function $(0 < p < \infty)$ in terms of functions in $H^{p_1}_{\mathrm{flag}}$ and $H^{p_2}_{\mathrm{flag}}$ with $0 < p_1 < p < p_2 < \infty$. This gives rise to an interpolation theorem between $H^{p_1}_{\mathrm{flag}}$ and $H^{p_2}_{\mathrm{flag}}$ for any $0 < p_2 < p_1 < \infty$ $(H^p_{\mathrm{flag}} = L^p$ for $1 < p < \infty)$.

We now describe our approach and results in more detail. Proofs will be given in subsequent parts of the paper.

2. The square function on the Heisenberg group

We begin with an *implicit* two-parameter continuous variant of the Littlewood–Paley square function that is introduced in [Müller et al. 1996]. For this we need the standard Calderón reproducing formula on the Heisenberg group. Note that spectral theory was used in place of the Calderón reproducing formula in [Müller et al. 1996].

Theorem 2 [Geller and Mayeli 2006, Corollary 1]. *There is $\psi \in C^\infty(\mathbb{H}^n)$ satisfying either*

$$\psi \in \mathscr{S}(\mathbb{H}^n) \text{ and all moments of } \psi \text{ vanish, or}$$

$$\psi \in C^\infty_c(\mathbb{H}^n) \text{ and all arbitrarily large moments of } \psi \text{ vanish,}$$

such that the following Calderón reproducing formula holds:

$$f = \int_0^\infty \psi_s^\vee * \psi_s * f \frac{ds}{s}, \quad f \in L^2(\mathbb{H}^n),$$

where $*$ is Heisenberg convolution, $\psi^\vee(\zeta) = \overline{\psi(\zeta^{-1})}$, and $\psi_s(z, t) = s^{-2n-2}\psi(z/s, u/s^2)$ for $s > 0$.

Remark 3. We will usually assume that ψ above has compact support. However, it will sometimes be convenient for us if the component functions ψ_s have *infinitely* many vanishing moments. In particular we can then use the *same* component functions to define the flag Hardy spaces for *all* $0 < p < \infty$ (the smaller p is, the more vanishing moments are required to obtain necessary decay of singular integrals). Thus we will sometimes sacrifice the property of having compactly supported component functions.

We now wish to extend this formula to encompass the flag structure on the Heisenberg group $\mathbb{H}^n$.

2.1. *The component functions.* Following [Müller et al. 1996], we construct a Littlewood–Paley *component* function ψ defined on $\mathbb{H}^n \simeq \mathbb{C}^n \times \mathbb{R}$, given by the partial convolution $*_2$ in the second variable only:

$$\psi(z, u) = \psi^{(1)} *_2 \psi^{(2)}(z, u) = \int_\mathbb{R} \psi^{(1)}(z, u - v)\psi^{(2)}(v)\, dv, \quad (z, u) \in \mathbb{C}^n \times \mathbb{R},$$

where $\psi^{(1)} \in \mathscr{S}(\mathbb{H}^n)$ is as in Theorem 2, and $\psi^{(2)} \in \mathscr{S}(\mathbb{R})$ satisfies

$$\int_0^\infty |\widehat{\psi^{(2)}}(t\eta)|^2 \frac{dt}{t} = 1$$

for all $\eta \in \mathbb{R} \backslash \{0\}$, along with the moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta \psi^{(1)}(z, u) \, dz \, du = 0, \quad |\alpha| + 2\beta \leq M,$$

$$\int_{\mathbb{R}} v^\gamma \psi^{(2)}(v) dv = 0, \quad \gamma \geq 0.$$

Here the positive integer M may be taken arbitrarily large when the support of $\psi^{(1)}$ is compact, and may be infinite otherwise.

Thus we have

$$f(z, u) = \int_0^\infty \int_0^\infty \check{\psi}_{s,t} * \psi_{s,t} * f(z, u) \frac{ds}{s} \frac{dt}{t}, \quad f \in L^2(\mathbb{H}^n), \tag{2-1}$$

where the functions $\psi_{s,t}$ are given by

$$\psi_{s,t}(z, u) = \psi_s^{(1)} *_2 \psi_t^{(2)}(z, u), \tag{2-2}$$

with

$$\psi_s^{(1)}(z, u) = s^{-2n-2} \psi^{(1)}\left(\frac{z}{s}, \frac{u}{s^2}\right) \quad \text{and} \quad \psi_t^{(2)}(v) = t^{-1} \psi^{(2)}\left(\frac{v}{t}\right),$$

and where the integrals in (2-1) converge in $L^2(\mathbb{H}^n)$. Indeed,

$$\check{\psi}_{s,t} *_{\mathbb{H}^n} \psi_{s,t} *_{\mathbb{H}^n} f(z, u) = (\check{\psi}_s^{(1)} *_2 \psi_t^{(2)}) *_{\mathbb{H}^n} (\check{\psi}_s^{(1)} *_2 \psi_t^{(2)}) *_{\mathbb{H}^n} f(z, u)$$

$$= (\check{\psi}_s^{(1)} *_{\mathbb{H}^n} \check{\psi}_s^{(1)}) *_{\mathbb{H}^n} (\psi_t^{(2)} *_{\mathbb{R}} \psi_t^{(2)}) *_2 f(z, u)$$

yields (2-1) upon invoking the standard Calderón reproducing formula on $\mathbb{R}$ and then Theorem 2 on $\mathbb{H}^n$:

$$\int_0^\infty \int_0^\infty \check{\psi}_{s,t} *_{\mathbb{H}^n} \psi_{s,t} *_{\mathbb{H}^n} f(z, u) \frac{ds}{s} \frac{dt}{t} = \int_0^\infty \check{\psi}_s^{(1)} *_{\mathbb{H}^n} \check{\psi}_s^{(1)} *_{\mathbb{H}^n} \left\{ \int_0^\infty \psi_t^{(2)} *_{\mathbb{R}} \psi_t^{(2)} *_2 f(z, u) \frac{dt}{t} \right\} \frac{ds}{s}$$

$$= \int_0^\infty \check{\psi}_s^{(1)} *_{\mathbb{H}^n} \check{\psi}_s^{(1)} *_{\mathbb{H}^n} f(z, u) \frac{ds}{s} = f(z, u).$$

For $f \in L^p$, $1 < p < \infty$, the *continuous* Littlewood–Paley square function $g_{\text{flag}}(f)$ of f is defined by

$$g_{\text{flag}}(f)(z, u) = \left\{ \int_0^\infty \int_0^\infty |\psi_{s,t} * f(z, u)|^2 \frac{ds}{s} \frac{dt}{t} \right\}^{\frac{1}{2}},$$

Note that we have the *flag* moment conditions, so called because they include only half of the product moment conditions associated with the product $\mathbb{C}^n \times \mathbb{R}$:

$$\int_{\mathbb{R}} u^\alpha \psi(z, u) \, du = 0 \quad \text{for all } \alpha \in \mathbb{Z}_+ \text{ and } z \in \mathbb{C}^n. \tag{2-3}$$

Indeed, with the change of variable $u' = u - v$ and the binomial theorem

$$(u' + v)^\beta = \sum_{\beta = \gamma + \delta} c_{\gamma, \delta} (u')^\gamma v^\delta,$$

we have

$$\int_{\mathbb{R}} u^{\alpha} \psi(z, u) \, du = \int_{\mathbb{R}} u^{\alpha} \left\{ \int_{\mathbb{R}} \psi^{(2)}(u - v) \psi^{(1)}(z, v) \, dv \right\} du$$

$$= \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} (u' + v)^{\alpha} \psi^{(2)}(u') \, du' \right\} \psi^{(1)}(z, v) \, dv$$

$$= \sum_{\alpha = \gamma + \delta} c_{\gamma, \delta} \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} (u')^{\gamma} \psi^{(2)}(u') \, du' \right\} v^{\delta} \psi^{(1)}(z, v) \, dv$$

$$= \sum_{\alpha = \gamma + \delta} c_{\gamma, \delta} \int_{\mathbb{R}} \{0\} v^{\delta} \psi^{(1)}(z, v) \, dv = 0$$

for all $\alpha \in \mathbb{Z}_+$ and each $z \in \mathbb{C}^n$. Note that, as a consequence, the *full* moments $\int_{\mathbb{H}^n} z^{\alpha} u^{\beta} \psi(z, u) \, dz \, du$ all vanish, but that, in general, the partial moments $\int_{\mathbb{C}^n} z^{\alpha} \psi(z, u) \, dz$ do *not* vanish.

Remark 4. As observed in [Nagel et al. 2012], there is a weak cancellation substitute for this failure to vanish, namely an *estimate* for $\int_{\mathbb{C}^n} z^{\alpha} \psi(z, u) \, dz$ that is derived from the vanishing moments of $\psi^{(1)}(z, v)$ and the smoothness of $\psi^{(2)}(u)$ via the identity

$$\int_{\mathbb{C}^n} z^{\alpha} \psi(z, u) \, dz = \int_{\mathbb{C}^n} \int_{\mathbb{R}} z^{\alpha} \psi^{(1)}(z, v) \psi^{(2)}(u - v) \, dz \, dv$$

$$= \int_{\mathbb{C}^n} \int_{\mathbb{R}} z^{\alpha} \psi^{(1)}(z, v) [\psi^{(2)}(u - v) - \psi^{(2)}(u)] \, dz \, dv.$$

We will not pursue this further here.

We will also consider the associated *sequence* of component functions $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$, where the functions $\psi_{j,k}$ are given by

$$\psi_{j,k}(z, u) = \psi_j^{(1)} *_2 \psi_k^{(2)}(z, u), \tag{2-4}$$

with

$$\psi_j^{(1)}(z, u) = 2^{\alpha j (2n+2)} \psi^{(1)}(2^{\alpha j} z, 2^{2\alpha j} u) \quad \text{and} \quad \psi_k^{(2)}(v) = 2^{2\alpha k} \psi^{(2)}(2^{2\alpha k} v),$$

and $\psi^{(1)}$ and $\psi^{(2)}$ as above. Here α is a small positive constant that will be fixed in Theorem 17 below, where we establish a *wavelet* Calderón reproducing formula using this sequence of component functions for small α. We then have a corresponding *discrete* (convolution) Littlewood–Paley square function $g_{\text{flag}}(f)$ defined by

$$g_{\text{flag}}(f)(z, u) = \left\{ \sum_j \sum_k |\psi_{j,k} * f(z, u)|^2 \right\}^{\frac{1}{2}}.$$

This should be compared with the analogous square function in [Müller et al. 1996].

Remark 5. The terminology "implicit two-parameter structure" is inspired by the fact that the functions $\psi_{s,t}(z, u)$ and $\psi_{j,k}(z, u)$ are not dilated directly from $\psi(z, u)$, but rather from a lifting of ψ to a product function. It is the subtle convolution $*_2$ that facilitates a passage from one-parameter "cubes" to two-parameter "rectangles" as dictated by the geometry of the kernels considered.

2.2. *Square function inequalities.* Altogether, we have from above that

$$f(z, u) = \int_0^\infty \int_0^\infty \psi_{s,t} * \psi_{s,t} * f(z, u) \frac{ds}{s} \frac{dt}{t}, \quad f \in L^2(\mathbb{H}^n). \tag{2-5}$$

Note that if one considers the integral on the right-hand side as an operator, then, by the construction of the function ψ, it is a flag singular integral operator and has the implicit multiparameter structure mentioned above. Using iteration and the vector-valued Littlewood–Paley estimate together with the Calderón reproducing formula on L^2 allows us to obtain L^p estimates for g_{flag}, $1 < p < \infty$, in Theorem 6 below. This should be compared to the variant in [Müller et al. 1996, Proposition 4.1] for g-functions constructed from spectral theory for $\mathcal{L}$ and T.

Theorem 6. *Let $1 < p < \infty$. There exist constants C_1 and C_2 depending on n and p such that*

$$C_1 \|f\|_p \leq \|g_{\text{flag}}(f)\|_p \leq C_2 \|f\|_p, \quad f \in L^p(\mathbb{H}^n).$$

In order to state our results for flag singular integrals on $\mathbb{H}^n$, we need to recall some definitions given in [Nagel et al. 2001]. We begin with the definition of a class of distributions on Euclidean space $\mathbb{R}^N$. A k-normalized bump function on a space $\mathbb{R}^N$ is a C^k-function supported on the unit ball with C^k norm bounded by 1. As pointed out in [Nagel et al. 2001], the definitions given below are independent of the choices of $k \geq 1$, and thus we will simply refer to a "normalized bump function" without specifying the index k.

We will rephrase Definition 2.1.1 in [Nagel et al. 2001] of a flag kernel in the case of the Heisenberg group as follows.

Definition 7. A flag convolution kernel on $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$ is a distribution K on $\mathbb{R}^{2n+1}$ which coincides with a C^∞ function away from the coordinate subspace $\{(0, u)\} \subset \mathbb{H}^n$, where $0 \in \mathbb{C}^n$ and $u \in \mathbb{R}$, and satisfies the following:

(1) (differential inequalities) For any multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\beta = (\beta_1, \ldots, \beta_m)$,

$$|\partial_z^\alpha \partial_u^\beta K(z, u)| \leq C_{\alpha,\beta} |z|^{-2n-|\alpha|} \cdot (|z|^2 + |u|)^{-1-|\beta|}$$

for all $(z, u) \in \mathbb{H}^n$ with $z \neq 0$.

(2) (cancellation condition) For every multi-index α and every normalized bump function ϕ_1 on $\mathbb{R}$ and every $\delta > 0$,

$$\left| \int_\mathbb{R} \partial_z^\alpha K(z, u) \phi_1(\delta u) \, du \right| \leq C_\alpha |z|^{-2n-|\alpha|};$$

for every multi-index β and every normalized bump function ϕ_2 on $\mathbb{C}^n$ and every $\delta > 0$,

$$\left| \int_{\mathbb{C}^n} \partial_u^\beta K(z, u) \phi_2(\delta z) \, dz \right| \leq C_\gamma |u|^{-1-|\beta|};$$

and for every normalized bump function ϕ_3 on $\mathbb{H}^n$ and every $\delta_1 > 0$ and $\delta_2 > 0$,

$$\left| \int_{\mathbb{H}^n} K(z, u) \phi_3(\delta_1 z, \delta_2 u) \, dz \, du \right| \leq C.$$

As in [Müller et al. 1995], we may always assume that a flag kernel $K(z, u)$ is integrable on $\mathbb{H}^n$ by using a smooth truncation argument.

Informally, we can now define the flag Hardy space $H^p_{\text{flag}}(\mathbb{H}^n)$ on the Heisenberg group for $0 < p \le 1$ by

$$H^p_{\text{flag}}(\mathbb{H}^n) = \{f \text{ a } \textit{distribution} \text{ on } \mathbb{H}^n : g_{\text{flag}}(f) \in L^p(\mathbb{H}^n)\},$$

and, for $f \in H^p_{\text{flag}}(\mathbb{H}^n)$, define

$$\|f\|_{H^p_{\text{flag}}} = \|g_{\text{flag}}(f)\|_p.$$

Of course we need to give a precise definition of *distribution* in this context, and a natural question then arises as to whether or not the resulting definition is independent of the choice of component functions $\psi_{j,k}$ in the definition of the square function g_{flag}. Moreover, to study the H^p_{flag}-boundedness of flag singular integrals and establish the duality theory of H^p_{flag}, this definition is difficult to use when $0 < p \le 1$. We need to approximately discretize the quasinorm of H^p_{flag}. In order to obtain this discrete H^p_{flag} quasinorm we will prove certain Plancherel–Pólya-type inequalities, and the main tool used in proving such inequalities will be the wavelet Calderón reproducing formula that we define below. To be more specific, we will prove that the formula (2-5) converges in certain spaces of test functions $\mathcal{M}^M_{\text{flag}}(\mathbb{H}^n)$ adapted to the flag structure, and thus also in the dual spaces $\mathcal{M}^M_{\text{flag}}(\mathbb{H}^n)'$ (see Theorem 17 below). Furthermore, using an approximation procedure and an almost-orthogonality argument, we prove in Theorem 17 below a *wavelet* Calderón reproducing formula which expresses f as a Fourier-like *series* of molecules or "wavelets" $(z, u) \to \tilde{\psi}_{j,k}(z, u, z_I, u_J)$ with coefficients $\psi_{j,k} * f(z_I, u_J)$.

In order to describe this formula explicitly in Section 3 below, we will use the *flag* dyadic decomposition

$$\mathbb{H}^n = \bigcup_{(\alpha,\tau)\in K_j}^{\cdot} \mathcal{S}_{j,\alpha,\tau}$$

of the Heisenberg group given in Theorem 68 below (this is a "hands on" variant of the tiling construction in [Strichartz 1992]), as well as the notion of Heisenberg rectangles

$$\mathcal{R}^{\mathcal{S}_{k,\beta,\upsilon}}_{\mathcal{S}_{j,\alpha,\tau}}(\text{ver}) \quad \text{and} \quad \mathcal{R}^{\mathcal{S}_{k,\beta,\upsilon}}_{\mathcal{S}_{j,\alpha,\tau}}(\text{hor})$$

given in Definition 69 below when $j \le k$ and $\mathcal{S}_{j,\alpha,\tau}$ and $\mathcal{S}_{k,\beta,\upsilon}$ are dyadic cubes in $\mathbb{H}^n$ with $\mathcal{S}_{j,\alpha,\tau} \subset \mathcal{S}_{k,\beta,\upsilon}$. Recall that

$$\{I\}_{I \text{ dyadic}} = \{I^j_\alpha\}_{j\in\mathbb{Z} \text{ and } \alpha\in 2^j\mathbb{Z}^{2n}}$$

is the usual dyadic grid in $\mathbb{C}^n$ and that

$$\{J\}_{J \text{ dyadic}} = \{J^k_\tau\}_{k\in\mathbb{Z} \text{ and } \tau\in 2^k\mathbb{Z}}$$

is the usual dyadic grid in $\mathbb{R}$. The projection of the dyadic cube $\mathcal{S}_{j,\alpha,\tau}$ onto $\mathbb{C}^n$ is the dyadic cube I^j_α, and

$$\mathcal{R}^{\mathcal{S}_{k,\beta,\upsilon}}_{\mathcal{S}_{j,\alpha,\tau}}(\text{ver}) \quad (\text{respectively } \mathcal{R}^{\mathcal{S}_{k,\beta,\upsilon}}_{\mathcal{S}_{j,\alpha,\tau}}(\text{hor}))$$

plays the role of the dyadic rectangle $I_\alpha^j \times J_\upsilon^{2k}$ (respectively $I_\beta^k \times J_\tau^{2j}$). In the Heisenberg group, these rectangles necessarily "rotate" with the group structure.

Notation 8. It will be convenient to use the suggestive, if somewhat imprecise, notation

$$\mathscr{R} = I \times J = I_\alpha^j \times J_\upsilon^{2k}$$

for the dyadic rectangle $\mathscr{R}_{\mathscr{S}_{j,\alpha,\tau}}^{\mathscr{S}_{k,\beta,\upsilon}}$ (ver), etc. It should be emphasized that $\mathscr{R} = I \times J$ is *not* a product set, but rather a dyadic Heisenberg rectangle $\mathscr{R}_{\mathscr{S}_{j,\alpha,\tau}}^{\mathscr{S}_{k,\beta,\upsilon}}$ (ver) that serves as a Heisenberg substitute for the actual product set I_α^j times J_υ^{2k}. Thus we will say that the dyadic rectangle $\mathscr{R} = I \times J$ has *side lengths* $\ell(I) = 2^j$ and $\ell(J) = 2^{2k}$. For $j \leq k$, the collection of all dyadic Heisenberg rectangles $\mathscr{R} = I \times J$ with side lengths 2^j and 2^{2k} will be denoted by

$$\mathscr{R}(2^j \times 2^{2k}) \equiv \left\{ \mathscr{R} = I \times J = I_\alpha^j \times J_\upsilon^{2k} = \mathscr{R}_{\mathscr{S}_{j,\alpha,\tau}}^{\mathscr{S}_{k,\beta,\upsilon}}(\text{ver}) : \mathscr{S}_{j,\alpha,\tau} \subset \mathscr{S}_{k,\beta,\upsilon} \right\}.$$

Caution: For $k \leq j$, the support of the component function $\psi_{j,k}$ defined in (2-4) is essentially a vertical Heisenberg rectangle $I \times J$ having side lengths $\ell(I) = 2^{-j}$ and $\ell(J) = 2^{-2k}$. Note the passage from j, k to $-j, -k$.

2.3. *Standard test functions.*

We now describe the features inherent in giving a precise definition of the flag Hardy space $H_{\text{flag}}^p(\mathbb{H}^n)$ as elements in the dual of familiar test spaces. We begin by introducing the test spaces $\mathscr{M}_{\text{flag}}^M(\mathbb{H}^n)$ associated with the flag structure on $\mathbb{H}^n$ that are obtained by *projecting* the corresponding product test spaces $\mathscr{M}_{\text{product}}^M(\mathbb{H}^n \times \mathbb{R})$ onto $\mathbb{H}^n$. Our definitions here will encompass the entire range $0 < p \leq 1$. For this we use the projection of functions F defined on $\mathbb{H}^n \times \mathbb{R}$ to functions $f = \pi F$ defined on $\mathbb{H}^n$ as introduced in [Müller et al. 1995]:

$$f(z, u) = (\pi F)(z, u) \equiv \int_{\mathbb{R}} F((z, u - v), v) \, dv. \tag{2-6}$$

We will also use the notation $\pi F = F_\flat$ as in [Müller et al. 1995]. Recall that $2n + 1$ is the Euclidean dimension of the Heisenberg group $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$ and that $Q = 2n + 2$ is the homogeneous dimension of $\mathbb{H}^n$.

In this notation, the component function $\psi(z, u)$ in Subsection 2.1 above is given by $\pi \Psi(z, u)$, where

$$\Psi(z, u, v) \equiv \psi^{(1)}(z, u) \psi^{(2)}(v). \tag{2-7}$$

We now define an appropriate *product* molecular space $\mathscr{M}_{\text{product}}^{M_1, M_2, M}$ on $\mathbb{H}^n \times \mathbb{R}$ with three parameters M_1, M_2, M.

Remark 9. Note that, in the definition below, we require *equally* many moments and derivatives in each of the u and v variables, and exactly *twice* as many moments and derivatives in the z variable. The integer M controls the decay of the function, the integer M_1 controls the total number of moments, and the integer M_2 controls the total weighted number of derivatives permitted.

Definition 10. Let $M, M_1, M_2 \in \mathbb{N}$ be positive integers and let $0 < \delta \leq 1$. The *product* molecular space $\mathscr{M}_{\text{product}}^{M+\delta, M_1, M_2}(\mathbb{H}^n \times \mathbb{R})$ consists of all functions $F((z, u), v)$ on $\mathbb{H}^n \times \mathbb{R}$ satisfying the product moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta F((z,u),v)\, dz\, du = 0 \quad \text{for all } |\alpha| + 2\beta \le M_1,$$
$$\int_{\mathbb{R}} v^\gamma F((z,u),v)\, dv = 0 \quad \text{for all } 2\gamma \le M_1, \tag{2-8}$$

and such that there is a nonnegative constant A satisfying the four differential inequalities

$$|\partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v)| \le A \frac{1}{(1+|z|^2+|u|)^{(Q+M+|\alpha|+2\beta+\delta)/2}} \frac{1}{(1+|v|)^{1+M+\gamma+\delta}}$$
$$\text{for all } |\alpha| + 2\beta \le M_2 \text{ and } 2\gamma \le M_2, \quad \text{(2-9)}$$

$$|\partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v) - \partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z',u'),v)|$$
$$\le A \frac{|(z,u)\circ(z',u')^{-1}|^\delta}{(1+|z|^2+|u|)^{(Q+M+M_2+2\delta)/2}} \frac{1}{(1+|v|)^{1+M+\gamma+\delta}}$$
$$\text{for all } |\alpha| + 2\beta = M_2 \text{ and } |(z,u)\circ(z',u')^{-1}| \le \tfrac{1}{2}(1+|z|^2+|u|)^{\frac{1}{2}}, \quad \text{(2-10)}$$

$$|\partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v) - \partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v')|$$
$$\le A \frac{1}{(1+|z|^2+|u|)^{(Q+M+|\alpha|+2\beta+\delta)/2}} \frac{|v-v'|^\delta}{(1+|v|)^{1+M+M_2/2+2\delta}},$$
$$\text{for all } |\alpha| + 2\beta \le M_2, \ 2\gamma = M_2, \text{ and } |v-v'| \le \tfrac{1}{2}(1+|v|), \quad \text{(2-11)}$$

$$\left|[\partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v) - \partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z',u'),v)] - [\partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z,u),v') - \partial_z^\alpha \partial_u^\beta \partial_v^\gamma F((z',u'),v')]\right|$$
$$\le A \frac{|(z,u)\circ(z',u')^{-1}|^\delta}{(1+|z|^2+|u|)^{(Q+M+M_2+2\delta)/2}} \frac{|v-v'|^\delta}{(1+|v|)^{1+M+M_2/2+2\delta}}$$

for all $|\alpha|+2\beta = M_2$, $2\gamma = M_2$, $|(z,u)\circ(z',u')^{-1}| \le \tfrac{1}{2}(1+|z|^2+|u|)^{\frac{1}{2}}$, and $|v-v'| \le \tfrac{1}{2}(1+|v|)$.　(2-12)

The space $\mathcal{M}_{\text{product}}^{M+\delta,M_1,M_2}(\mathbb{H}^n \times \mathbb{R})$ becomes a Banach space under the norm defined by the least nonnegative number A for which the above four inequalities hold.

Now we define the *flag* molecular space $\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)$ as the projection of $\mathcal{M}_{\text{product}}^{M+\delta,M_1,M_2}(\mathbb{H}^n \times \mathbb{R})$ under the map π given in (2-6).

Definition 11. Let $M, M_1, M_2 \in \mathbb{N}$ be positive integers and $0 < \delta \le 1$. The *flag* molecular space $\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)$ consists of all functions f on $\mathbb{H}^n$ such that there is $F \in \mathcal{M}_{\text{product}}^{M+\delta,M_1,M_2}(\mathbb{H}^n \times \mathbb{R})$ with $f = \pi F = F_\flat$. Define a norm on $\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)$ by

$$\|f\|_{\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)} \equiv \inf_{F:\, f=\pi F} \|F\|_{\mathcal{M}_{\text{product}}^{M+\delta,M_1,M_2}(\mathbb{H}^n \times \mathbb{R})}.$$

Thus the norm on $\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)$ is the quotient norm

$$\|f\|_{\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)} = \mathcal{M}_{\text{product}}^{M+\delta,M_1,M_2}(\mathbb{H}^n \times \mathbb{R}) / \pi^{-1}(\{0\}),$$

and $\mathcal{M}_{\text{flag}}^{M+\delta,M_1,M_2}(\mathbb{H}^n)$ is a Banach space.

We record here an intertwining formula for π and a convolution operator T on $\mathbb{H}^n$. Let

$$Tf(z, u) = K *_{\mathbb{H}^n} f(z, u) = \int_{\mathbb{H}^n} K((z, u) \circ (z', u')^{-1}) f(z', u')\, dz'\, du'.$$

Extend T to an operator $\widetilde{T} = T \otimes \delta_0$ on the group $\mathbb{H}^n \times \mathbb{R}$ by acting T in the $\mathbb{H}^n$ factor only:

$$\widetilde{T} F((z, u), v) = \int_{\mathbb{H}^n} K((z, u) \circ (z', u')^{-1}) F(z', u', v)\, dz'\, du'.$$

Lemma 12. *Let T be a convolution operator on $\mathbb{H}^n$ and let $\widetilde{T} = T \otimes \delta_0$ be its extension to $\mathbb{H}^n \times \mathbb{R}$ defined above. Then*

$$T(\pi F)(z, u) = \pi(\widetilde{T} F)(z, u).$$

Proof. Formally we have

$$\begin{aligned}
T(\pi F)(z, u) &= \int_{\mathbb{H}^n} K((z, u) \circ (z', u')^{-1})(\pi F)(z', u')\, dz'\, du' \\
&= \int_{\mathbb{H}^n} K((z, u) \circ (z', u')^{-1}) \left\{ \int_{\mathbb{R}} F(z', u' - v, v)\, dv \right\} dz'\, du' \\
&= \int_{\mathbb{H}^n} \int_{\mathbb{R}} K(z - z', u - u' + 2\operatorname{Im} \overline{z'}z) F(z', u' - v, v)\, dv\, dz'\, du'.
\end{aligned}$$

Now make the change of variable $w' = u' - v$ to get

$$\begin{aligned}
T(\pi F)(z, u) &= \int_{\mathbb{H}^n} \int_{\mathbb{R}} K(z - z', u - w' - v + 2\operatorname{Im} \overline{z'}z) F(z', w', v)\, dv\, dz'\, dw' \\
&= \int_{\mathbb{R}} \left\{ \int_{\mathbb{H}^n} K((z, u - v) \circ (z', w')^{-1}) F(z', w', v)\, dz'\, dw' \right\} dv \\
&= \int_{\mathbb{R}} \{ \widetilde{T} F(z, u - v, v) \}\, dv = \pi(\widetilde{T} F)(z, u). \qquad \square
\end{aligned}$$

Later in the paper we will fix $M_1 = M_2 = M$ and denote $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ simply by $\mathscr{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)$, but for now we will allow M_1 and M_2 to remain independent of M in order to further analyze the space $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$.

2.3.1. *An analysis of the projected flag molecular space.* Lemma 14 below shows that functions $f(z, u)$ in the "projected" flag molecular space $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ have moments in the u variable alone, as well as *more* moments in the (z, u) variable than we might expect. We refer loosely to this situation as having *half-product* moments. There is a more familiar space of test functions $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$, defined below with half-product moments, that *avoids* the operation of projection, and that is closely related to the projected test space $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$. While we do not know if the spaces $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ and $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ coincide, the embeddings in Lemma 14 below are enough for our purposes.

Definition 13. Let M, M_1, $M_2 \in \mathbb{N}$ be positive integers and $0 < \delta \leq 1$. Define the "moment" *flag* molecular space $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ to consist of all functions f on $\mathbb{H}^n$ satisfying the moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta f(z, u)\, dz\, du = 0 \quad \text{for all } |\alpha| \leq M_1,\ |\alpha| + 2\beta \leq 2M_1 + 2,$$

$$\int_{\mathbb{R}} u^\gamma f(z, u)\, du = 0 \quad \text{for all } \gamma \leq M_1,$$

and such that there is a nonnegative constant A satisfying the differential inequalities

$$|\partial_z^\alpha \partial_u^\beta f(z, u)| \leq A \frac{1}{(1 + |z|^2 + |u|)^{(Q+M+|\alpha|+2\beta)/2}} \quad \text{for all } |\alpha| + 2\beta \leq M_2,$$

$$|\partial_z^\alpha \partial_u^\beta f(z, u) - \partial_z^\alpha \partial_u^\beta f(z', u')| \leq A \frac{|(z, u) \circ (z', u')^{-1}|^\delta}{(1 + |z|^2 + |u|)^{(Q+M+\delta+M_2)/2}}$$

$$\text{for all } |\alpha| + 2\beta = M_2 \text{ and } |(z, u) \circ (z', u')^{-1}| \leq \tfrac{1}{2}(1 + |z|^2 + |u|)^{\frac{1}{2}}.$$

Note that the moment conditions in the definition of $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ permit larger values of β depending on $|\alpha|$ than in the definition of $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$. The space $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ becomes a Banach space under the norm defined by the least nonnegative number A for which the above two inequalities hold.

Lemma 14. *The spaces* $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ *and* $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ *satisfy the containments*

$$\mathsf{M}_F^{3M+\delta+M_2, M_1, 2M_2+4}(\mathbb{H}^n) \subset \mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n) \subset \mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n),$$

which are continuous:

$$\|f\|_{\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)} \lesssim \|f\|_{\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)} \lesssim \|f\|_{\mathsf{M}_F^{3M+\delta+M_2, M_1, 2M_2+4}(\mathbb{H}^n)}.$$

Remark 15. The importance of the "projected" flag molecular space $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ lies in the existence of a *wavelet* Calderón reproducing formula for this space of test functions; see Theorem 17 below. We do not know if such a reproducing formula holds for the "moment" flag space $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$, but the embeddings in Lemma 14 will prove important in identifying the distributions in the dual space $\mathscr{M}_{\text{flag}}^{M+\delta, M_1, M_2}(\mathbb{H}^n)'$ as being "roughly" those in a dual space $\mathsf{M}_F^{M'+\delta, M_1', M_2'}(\mathbb{H}^n)'$.

Remark 16. The integer M_1 that controls the number of moments in $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$ remains the same in both the smaller space $\mathsf{M}_F^{3M+\delta+M_2, M_1, 2M_2+4}(\mathbb{H}^n)$ and the larger space $\mathsf{M}_F^{M+\delta, M_1, M_2}(\mathbb{H}^n)$. However, we lose both derivatives and decay in passing from the smaller to the larger space.

While we cannot say that $H_{\text{flag}}^p(\mathbb{H}^n)$ is a *subspace* of the more familiar one-parameter Hardy space $H^p(\mathbb{H}^n)$, we can show that the quotient space

$$Q_{\text{flag}}^p(\mathbb{H}^n) \equiv H_{\text{flag}}^p(\mathbb{H}^n) / \mathsf{M}_F^{M'+\delta, M_1', M_2'}(\mathbb{H}^n)^\perp$$

of $H_{\text{flag}}^p(\mathbb{H}^n)$ can be identified with a closed subspace of the corresponding quotient space

$$Q^p(\mathbb{H}^n) \equiv H^p(\mathbb{H}^n) / \mathsf{M}_F^{M'+\delta, M_1', M_2'}(\mathbb{H}^n)^\perp$$

of $H^p(\mathbb{H}^n)$, thus giving a sense in which the distributions we use to define $H_{\text{flag}}^p(\mathbb{H}^n)$ are "roughly" the same as those used to define $H^p(\mathbb{H}^n)$. See [Han et al. 2012] for details.

3. The wavelet Calderón reproducing formula

We can now state our *wavelet* Calderón reproducing formula for the flag structure in terms of the projected product test spaces

$$\mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n) \equiv \mathcal{M}_{\text{flag}}^{M+\delta, M.M}(\mathbb{H}^n),$$

defined by projecting the product test spaces

$$\mathcal{M}_{\text{product}}^{M+\delta}(\mathbb{H}^n \times \mathbb{R}) \equiv \mathcal{M}_{\text{product}}^{M+\delta, M, M}(\mathbb{H}^n \times \mathbb{R}).$$

We remind the reader that Euclidean versions of such reproducing formulas were obtained by Frazier and Jawerth [1990] using the Fourier transform together with the very special property that $\mathbb{R}^n$ is *tiled* by the compact abelian torus $\mathbb{T}^n$ and its discrete dual group, the lattice $\mathbb{Z}^n$.

It is convenient to introduce some new notation for the dyadic rectangles defined in Notation 8. Given $0 < \alpha < 1$ and a positive integer N, we write

$$\mathsf{R}(j, k) \equiv \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(k+N)}),$$
$$\mathsf{Q}(j) \equiv \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(j+N)}).$$

Now, for $\mathcal{Q} \in \mathsf{Q}(j)$, let $(z_{\mathcal{Q}}, u_{\mathcal{Q}})$ be any *fixed* point in the cube $\mathcal{Q}$, and for $\mathcal{R} \in \mathsf{R}(j, k)$ with $k < j$, let $(z_{\mathcal{R}}, u_{\mathcal{R}})$ be any *fixed* point in the rectangle $\mathcal{R}$. Let us write the collection of *all* dyadic cubes as

$$\mathsf{Q} \equiv \bigcup_{j \in \mathbb{Z}} \mathsf{Q}(j),$$

and the collection of *all strictly vertical* dyadic rectangles as

$$\mathsf{R}_{\text{vert}} \equiv \bigcup_{j > k} \mathsf{R}(j, k).$$

We now set

$$\psi'_{\mathcal{Q}} = \psi_j^{(1)} \qquad\qquad \text{if } \mathcal{Q} \in \mathsf{Q}(j),$$
$$\psi'_{\mathcal{R}} = \psi_{j,k} = \psi_j^{(1)} *_2 \psi_k^{(2)} \quad \text{if } \mathcal{R} \in \mathsf{R}(j, k),$$

where the $\psi_{j,k}$ are as in (2-4). Given an appropriate distribution f on $\mathbb{H}^n$, we define its *wavelet coefficients* $f_{\mathcal{Q}}$ and $f_{\mathcal{R}}$ by

$$f_{\mathcal{Q}} = \psi'_{\mathcal{Q}} * f(z_{\mathcal{Q}}, u_{\mathcal{Q}}) \quad \text{if } \mathcal{Q} \in \mathsf{Q},$$
$$f_{\mathcal{R}} = \psi'_{\mathcal{R}} * f(z_{\mathcal{R}}, u_{\mathcal{R}}) \quad \text{if } \mathcal{R} \in \mathsf{R}_{\text{vert}}, \text{ that is, when } j > k.$$

Below is the *wavelet* Calderón reproducing formula.

Theorem 17. *Suppose the notation is as above. Then there are associated functions $\tilde{\psi}_{\mathcal{Q}}, \tilde{\psi}_{\mathcal{R}} \in \mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)$ for $\mathcal{Q} \in \mathsf{Q}$ and $\mathcal{R} \in \mathsf{R}_{\text{vert}}$ satisfying*

$$\|\tilde{\psi}_{\mathcal{Q}}\|_{\mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)} \lesssim \|\psi'_{\mathcal{Q}}\|_{\mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)}, \quad \mathcal{Q} \in \mathsf{Q},$$

$$\|\tilde{\psi}_{\mathcal{R}}\|_{\mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)} \lesssim \|\psi'_{\mathcal{R}}\|_{\mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)}, \quad \mathcal{R} \in \mathsf{R}_{\text{vert}},$$

and

$$f(z, u) = \sum_{\mathcal{Q} \in \mathsf{Q}} f_{\mathcal{Q}} \tilde{\psi}_{\mathcal{Q}}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} f_{\mathcal{R}} \tilde{\psi}_{\mathcal{R}}(z, u), \quad (z, u) \in \mathbb{H}^n, \tag{3-1}$$

where the series in (3-1) converges in three spaces:

(1) *in $L^p(\mathbb{H}^n)$ for $1 < p < \infty$,*

(2) *in the Banach space $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$ for M' large enough,*

(3) *and in the corresponding dual space $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)'$ for M' large enough.*

Remark 18. Note that only *half* of the collection of dyadic rectangles, namely the vertical ones R_{vert}, are used in the wavelet Calderón reproducing formula. This is a reflection of the implicit product structure inherent in the Heisenberg group $\mathbb{H}^n$.

3.1. *Plancherel–Pólya inequalities and flag Hardy spaces.* The wavelet Calderón reproducing formula (3-1) yields the following Plancherel–Pólya-type inequalities; cf. [Pólya 1936; Plancherel and Pólya 1937]. We use the notation $A \approx B$ to indicate that two quantities A and B are comparable.

Theorem 19. *Suppose $\psi^{(1)}, \phi^{(1)} \in \mathcal{S}(\mathbb{C}^n)$ and $\psi^{(2)}, \phi^{(2)} \in \mathcal{S}(\mathbb{R})$, and let*

$$\psi(z, u) = \int_{\mathbb{R}} \psi^{(1)}(z, u - v) \psi^{(2)}(v) \, dv,$$

$$\phi(z, u) = \int_{\mathbb{R}} \phi^{(1)}(z, u - v) \psi^{(2)}(v) \, dv$$

be two component functions that each satisfies the conditions in Section 2.1. Then with Q, R_{vert}, $\psi'_{\mathcal{Q}}$, and $\psi'_{\mathcal{R}}$ as above, and for $f \in \mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)'$, $0 < p < \infty$, and M chosen large enough depending on n and p,

$$\left\| \left\{ \sum_{\mathcal{Q} \in \mathsf{Q}} \sup_{(z', u') \in \mathcal{Q}} |\psi'_{\mathcal{Q}} * f(z', u')|^2 \chi_{\mathcal{Q}}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} \sup_{(z', u') \in \mathcal{R}} |\psi'_{\mathcal{R}} * f(z', u')|^2 \chi_{\mathcal{R}}(z, u) \right\}^{\frac{1}{2}} \right\|_{L^p(\mathbb{H}^n)}$$

$$\approx \left\| \left\{ \sum_{\mathcal{Q} \in \mathsf{Q}} \inf_{(z', u') \in \mathcal{Q}} |\psi'_{\mathcal{Q}} * f(z', u')|^2 \chi_{\mathcal{Q}}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} \inf_{(z', u') \in \mathcal{R}} |\psi'_{\mathcal{R}} * f(z', u')|^2 \chi_{\mathcal{R}}(z, u) \right\}^{\frac{1}{2}} \right\|_{L^p(\mathbb{H}^n)}.$$

The Plancherel–Pólya-type inequalities in Theorem 19 will prove useful in establishing properties of the *wavelet* Littlewood–Paley g-function

$$g_{\text{flag}}(f)(z, u) = \left\{ \sum_{\mathcal{Q} \in \mathsf{Q}} |\psi'_{\mathcal{Q}} * f(z_{\mathcal{Q}}, u_{\mathcal{Q}})|^2 \chi_{\mathcal{Q}}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} |\psi'_{\mathcal{R}} * f(z_{\mathcal{R}}, u_{\mathcal{R}})|^2 \chi_{\mathcal{R}}(z, u) \right\}^{\frac{1}{2}},$$

where we are using the notation of Theorems 17 and 19.

We can now give a precise definition of the *flag* Hardy spaces.

Definition 20. Let $0 < p < \infty$. Then, for M sufficiently large depending on n and p, we define the *flag* Hardy space $H^p_{\text{flag}}(\mathbb{H}^n)$ on the Heisenberg group by

$$H^p_{\text{flag}}(\mathbb{H}^n) = \{ f \in \mathcal{M}^{M+\delta}_{\text{flag}}(\mathbb{H}^n)' : g_{\text{flag}}(f) \in L^p(\mathbb{H}^n) \},$$

and, for $f \in H^p_{\text{flag}}(\mathbb{H}^n)$, we set

$$\| f \|_{H^p_{\text{flag}}} = \| g_{\text{flag}}(f) \|_p. \tag{3-2}$$

Remark 21. We can take M in Definition 20 to satisfy

$$M \geq M_{n,p} \equiv (2n+2)\left[\frac{2}{p} - 1\right] + 1.$$

We have not computed the optimal value of $M_{n,p}$.

It is easy to see using Theorem 19 that the Hardy space H^p_{flag} in Definition 20 is well defined and that the H^p_{flag} norm of f is equivalent to the L^p norm of g_{flag}. By use of the Plancherel–Pólya-type inequalities, we will prove the boundedness of flag singular integrals on H^p_{flag} below.

3.2. *Boundedness of singular integrals and Marcinkiewicz multipliers.* Our main result is the $H^p_{\text{flag}} \to H^p_{\text{flag}}$ boundedness of flag singular integrals.

Theorem 22. *Suppose that T is a flag singular integral with kernel $K(z, u)$ as in* Definition 7. *Then T is bounded on H^p_{flag} for $0 < p \leq 1$. Namely, for all $0 < p \leq 1$ there exists a constant $C_{p,n}$ such that*

$$\| Tf \|_{H^p_{\text{flag}}} \leq C_{p,n} \| f \|_{H^p_{\text{flag}}}.$$

To obtain the $H^p_{\text{flag}} \to L^p$ boundedness of flag singular integrals, we prove the following general result:

Theorem 23. *Let $0 < p \leq 1$. If T is a linear operator which is bounded simultaneously on $L^2(\mathbb{R}^{2n+1})$ and $H^p_{\text{flag}}(\mathbb{H}^n)$, then T can be extended to a bounded operator from $H^p_{\text{flag}}(\mathbb{H}^n)$ to $L^p(\mathbb{R}^{2n+1})$.*

Remark 24. From the proof given in the next part of the paper, we see that this result holds in a larger setting, which includes the classical one-parameter and product Hardy spaces and the Hardy spaces on spaces of homogeneous type. Thus this provides an alternative approach to using Fefferman's criterion on boundedness of a singular integral operator by restricting its action on rectangle atoms [Fefferman 1986], and then combining this with Journé's geometric lemma; see [Journé 1985; 1986; Pipher 1986].

In particular, for flag singular integrals we can deduce the following.

Corollary 25. *Let T be a flag singular integral as in* Theorem 23. *Then T is bounded from $H^p_{\text{flag}}(\mathbb{H}^n)$ to $L^p(\mathbb{R}^{2n+1})$ for $0 < p \leq 1$.*

Remark 26. The conclusions of both Theorem 22 and Corollary 25 persist if we only require the moment and smoothness conditions on the flag kernel in Definition 7 to hold for $|\alpha|, \beta \leq N_{n,p}$, where $N_{n,p} < \infty$ is taken sufficiently large.

As a consequence, we can extend the Marcinkiewicz multiplier theorem in [Müller et al. 1995] (see Lemma 2.1 there) to flag Hardy spaces for $0 < p \leq 1$. To describe this extension, recall the standard sub-Laplacian $\mathscr{L}$ on the Heisenberg group

$$\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R} = \{(z,t) : z = (z_j)_{j=1}^n,\ z_j = x_j + iy_j \in \mathbb{C},\ t \in \mathbb{R}\},$$

defined by

$$\mathscr{L} \equiv -\sum_{j=1}^n (X_j^2 + Y_j^2), \quad X_j = \frac{\partial}{\partial x_j} + 2y_j \frac{\partial}{\partial t}, \quad Y_j = \frac{\partial}{\partial y_j} - 2x_j \frac{\partial}{\partial t}.$$

The operators $\mathscr{L}$ and $T = \partial/(\partial t)$ commute, and so do their spectral measures $dE_1(\xi)$ and $dE_2(\eta)$. Given a bounded function $m(\xi, \eta)$ on $\mathbb{R}_+ \times \mathbb{R}$, define the multiplier operator $m(\mathscr{L}, iT)$ on $L^2(\mathbb{H}^n)$ by

$$m(\mathscr{L}, iT) = \iint_{\mathbb{R}_+ \times \mathbb{R}} m(\xi, \eta)\, dE_1(\xi)\, dE_2(\eta).$$

Then $m(\mathscr{L}, iT)$ is automatically bounded on $L^2(\mathbb{H}^n)$, and if we impose Marcinkiewicz conditions on the multiplier, we obtain boundedness on flag Hardy spaces; this despite the fact that m is invariant under a two-parameter family of dilations $\delta_{(s,t)}$ which are group automorphisms only when $t = s^2$.

Corollary 27. *Let $0 < p \leq 1$, and suppose that $m(\xi, \eta)$ is a bounded function defined on $\mathbb{R}_+ \times \mathbb{R}$ that satisfies the Marcinkiewicz conditions*

$$|(\xi \partial_\xi)^\alpha (\eta \partial_\eta)^\beta m(\xi, \eta)| \leq C_{\alpha, \beta}$$

for all $|\alpha|, \beta \leq N_{n,p}$, where $N_{n,p} < \infty$ is taken sufficiently large. Then $m(\mathscr{L}, iT)$ is a bounded operator on $H_{\mathrm{flag}}^p(\mathbb{H}^n)$ for $0 < p \leq 1$.

The corollary follows from the results above together with [Müller et al. 1995, Theorem 3.1], which shows that the kernel $K(z, u)$ of a Marcinkiewicz multiplier $m(\mathscr{L}, iT)$ satisfies the conditions defining a flag convolution kernel in Definition 7.

3.3. *Carleson measures and duality.* To study the dual space of H_{flag}^p, we introduce the Carleson measure space $\mathrm{CMO}_{\mathrm{flag}}^p$.

Notation 28. It will often be convenient from now on to bundle the set Q of all dyadic cubes and the set $\mathsf{R}_{\mathrm{vert}}$ of all vertical dyadic rectangles into a single set

$$\mathsf{R}_+ = \mathsf{Q} \cup \mathsf{R}_{\mathrm{vert}}$$

consisting of all dyadic cubes and all vertical dyadic rectangles. We also write

$$\psi_{\mathscr{R}} = \begin{cases} \psi'_{\mathscr{Q}} & \text{if } \mathscr{R} = \mathscr{Q} \in \mathsf{Q}, \\ \psi'_{\mathscr{R}} & \text{if } \mathscr{R} \in \mathsf{R}_{\mathrm{vert}}. \end{cases}$$

Definition 29. Let $\psi_{j,k}$ be as in (2-4) with notation as above. We say that $f \in \mathrm{CMO}_{\mathrm{flag}}^p$ if $f \in \mathcal{M}_{\mathrm{flag}}^{M+\delta}(\mathbb{H}^n)'$ and the norm $\|f\|_{\mathrm{CMO}_{\mathrm{flag}}^p}$ is finite, where

$$\|f\|_{\mathrm{CMO}_{\mathrm{flag}}^p} \equiv \sup_{\Omega} \left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_{\mathcal{R} \in \mathrm{R}_+} \int_{\Omega} \sum_{\mathcal{R} \subset \Omega} |\psi_{\mathcal{R}} * f(z,u)|^2 \chi_{\mathcal{R}}(z,u) \, dz \, du \right\}^{\frac{1}{2}}$$

for all open sets Ω in $\mathbb{H}^n$ with finite measure.

Note that the Carleson measure condition is used with the implicit multiparameter structure in $\mathrm{CMO}_{\mathrm{flag}}^p$. When $p = 1$, we denote the space $\mathrm{CMO}_{\mathrm{flag}}^1$ as usual by $\mathrm{BMO}_{\mathrm{flag}}$. To see that the space $\mathrm{CMO}_{\mathrm{flag}}^p$ is well defined, one needs to show that the definition of $\mathrm{CMO}_{\mathrm{flag}}^p$ is independent of the choice of the component functions $\psi_{j,k}$. This can be proved just as for the Hardy space H_{flag}^p, using the following Plancherel–Pólya-type inequality.

Theorem 30. *Suppose ψ, ϕ satisfy the conditions as in Theorem 19. Then, for $f \in \mathcal{M}_{\mathrm{flag}}^{M+\delta}(\mathbb{H}^n)'$,*

$$\sup_{\Omega} \left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_{\mathcal{R} \in \mathrm{R}_+} \sum_{\mathcal{R} \subset \Omega} \sup_{(z,u) \in \mathcal{R}} |\psi_{\mathcal{R}} * f(z,u)|^2 |\mathcal{R}| \right\}^{\frac{1}{2}} \approx \sup_{\Omega} \left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_{\mathcal{R} \in \mathrm{R}_+} \sum_{\mathcal{R} \subset \Omega} \inf_{(z,u) \in \mathcal{R}} |\psi_{\mathcal{R}} * f(z,u)|^2 |\mathcal{R}| \right\}^{\frac{1}{2}},$$

where Ω ranges over all open sets in $\mathbb{H}^n$ with finite measure.

To show that $\mathrm{CMO}_{\mathrm{flag}}^p$ is the dual of H_{flag}^p, we introduce appropriate sequence spaces.

Definition 31. Let s^p be the collection of all sequences $s = \{s_{\mathcal{R}}\}_{\mathcal{R} \in \mathrm{R}_+}$ such that

$$\|s\|_{s^p} = \left\| \left\{ \sum_{\mathcal{R} \in \mathrm{R}_+} |s_{\mathcal{R}}|^2 |\mathcal{R}|^{-1} \chi_{\mathcal{R}} \right\}^{\frac{1}{2}} \right\|_{L^p(\mathbb{H}^n)} < \infty.$$

Let c^p be the collection of all sequences $s = \{s_{\mathcal{R}}\}$ such that

$$\|s\|_{c^p} = \sup_{\Omega} \left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_{\mathcal{R} \in \mathrm{R}_+} \sum_{\mathcal{R} \subset \Omega} |s_{\mathcal{R}}|^2 \right\}^{\frac{1}{2}} < \infty,$$

where Ω ranges over all open sets in $\mathbb{H}^n$ with finite measure.

We point out that only certain of the dyadic rectangles are used in s^p and c^p and these choices reflect the implicit multiparameter structure. Moreover, the Carleson measure condition is used in the definition of c^p. Next, we obtain the following duality theorem for sequence spaces.

Theorem 32. *Let $0 < p \leq 1$. Then we have $(s^p)^* = c^p$. More precisely, the map which sends $s = \{s_{\mathcal{R}}\}$ to $\langle s, t \rangle \equiv \sum_{\mathcal{R}} s_{\mathcal{R}} \bar{t}_{\mathcal{R}}$ defines a continuous linear functional on s^p with operator norm $\|t\|_{(s^p)^*} \approx \|t\|_{c^p}$, and, moreover, every $\ell \in (s^p)^*$ is of this form for some $t \in c^p$.*

When $p = 1$, this theorem in the one-parameter setting on $\mathbb{R}^n$ was proved in [Frazier and Jawerth 1990]. The proof given in [Frazier and Jawerth 1990] depends on estimates of certain distribution functions, which seem to be difficult to apply to the multiparameter case. For all $0 < p \leq 1$, we give a simple and more constructive proof of Theorem 32, which uses a stopping time argument for sequence spaces.

Theorem 32 together with the discrete Calderón reproducing formula and the Plancherel–Pólya-type inequalities yield the duality of H_{flag}^p.

Theorem 33. *Let $0 < p \leq 1$. Then*

$$(H_{\text{flag}}^p)^* = \text{CMO}_F^p.$$

More precisely, if $g \in \text{CMO}_{\text{flag}}^p$, the map ℓ_g given by $\ell_g(f) = \langle f, g \rangle$, defined initially for $f \in \mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)$, extends to a continuous linear functional on H_{flag}^p with $\|\ell_g\| \approx \|g\|_{\text{CMO}_{\text{flag}}^p}$. Conversely, for every $\ell \in (H_{\text{flag}}^p)^*$, there exists some $g \in \text{CMO}_{\text{flag}}^p$ so that $\ell = \ell_g$. In particular, $(H_{\text{flag}}^1)^* = \text{BMO}_{\text{flag}}$.

As a consequence of the duality of H_{flag}^1 and BMO_{flag}, together with the H_{flag}^1-boundedness of flag singular integrals, we obtain the BMO_{flag}-boundedness of flag singular integrals. Furthermore, we will see that $L^\infty \subseteq \text{BMO}_{\text{flag}}$ and hence the $L^\infty \to \text{BMO}_{\text{flag}}$ boundedness of flag singular integrals is also obtained. These provide the endpoint results of [Müller et al. 1995; Nagel et al. 2001], and can be summarized as follows.

Theorem 34. *Suppose that T is a flag singular integral with kernel as in Definition 7. Then T is bounded on BMO_{flag}. Moreover, there exists a constant C such that*

$$\|T(f)\|_{\text{BMO}_{\text{flag}}} \leq C \|f\|_{\text{BMO}_{\text{flag}}}.$$

3.4. *Calderón–Zygmund decompositions and interpolation.* Now we give the Calderón–Zygmund decomposition and interpolation theorems for flag Hardy spaces. We note that $H_{\text{flag}}^p(\mathbb{H}^n) = L^p(\mathbb{R}^{2n+1})$ for $1 < p < \infty$ by Theorem 6.

Theorem 35 (Calderón–Zygmund decomposition for flag Hardy spaces). *Let $0 < p_2 \leq 1$, $p_2 < p < p_1 < \infty$, let $\alpha > 0$ be given, and suppose $f \in H_{\text{flag}}^p(\mathbb{H}^n)$. Then we can write*

$$f = g + b,$$

where $g \in H_{\text{flag}}^{p_1}(\mathbb{H}^n)$ with $p < p_1 < \infty$ and $b \in H_{\text{flag}}^{p_2}(\mathbb{H}^n)$ with $0 < p_2 < p$, such that

$$\|g\|_{H_{\text{flag}}^{p_1}}^{p_1} \leq C\alpha^{p_1-p} \|f\|_{H_{\text{flag}}^p}^p \quad \text{and} \quad \|b\|_{H_{\text{flag}}^{p_2}}^{p_2} \leq C\alpha^{p_2-p} \|f\|_{H_{\text{flag}}^p}^p,$$

where C is an absolute constant.

Theorem 36 (interpolation theorem on flag Hardy spaces). *Let $0 < p_2 < p_1 < \infty$ and let T be a linear operator which is bounded from $H_{\text{flag}}^{p_2}$ to L^{p_2} and bounded from $H_{\text{flag}}^{p_1}$ to L^{p_1}. Then T is bounded from H_{flag}^p to L^p for all $p_2 < p < p_1$. Similarly, if T is bounded on $H_{\text{flag}}^{p_2}$ and $H_{\text{flag}}^{p_1}$, then T is bounded on H_{flag}^p for all $p_2 < p < p_1$.*

Remark 37. Combining Theorem 36 with Corollary 27 recovers the L^p boundedness of Marcinkiewicz multipliers in [Müller et al. 1995] (but not the sharp versions in [Müller et al. 1996]).

We point out that the Calderón–Zygmund decomposition in pure product domains for all L^p functions ($1 < p < 2$) into H^1 and L^2 functions, as well as the corresponding interpolation theorem, was established by Chang and Fefferman [1985; 1982].

Part II. Proofs of results

Part II of this paper contains the proofs of the results stated in Part I, and is organized as follows.

(1) In Section 4, we establish L^p estimates for the multiparameter Littlewood–Paley g-function when $1 < p < \infty$, and prove Theorems 6 and 38.

(2) In Section 5, we show that the Calderón reproducing formula holds on the flag molecular test function space $\mathcal{M}_{\mathrm{flag}}^{M+\delta}$ and its dual space $(\mathcal{M}_{\mathrm{flag}}^{M+\delta})'$. Then we prove the almost-orthogonality estimates and establish the wavelet Calderón reproducing formula on $\mathcal{M}_{\mathrm{flag}}^{M+\delta}$ and $(\mathcal{M}_{\mathrm{flag}}^{M+\delta})'$ in Theorem 17. Some estimates are established for the strong maximal function, and together with the wavelet Calderón reproducing formula, we then derive the Plancherel–Pólya-type inequalities in Theorem 19.

(3) In Section 6, we give a general result for bounding the L^p norm of the function by its H_{flag}^p norm (Theorem 56). We then prove the H_{flag}^p boundedness of flag singular integrals for all $0 < p \le 1$ in Theorem 22. The boundedness from H_{flag}^p to L^p for all $0 < p \le 1$ for the flag singular integral operators, Theorem 23, is thus a consequence of Theorem 22 and Theorem 56.

(4) Duality theory for the Hardy space H_{flag}^p is then established in Section 7 along with the boundedness of flag singular integral operators on $\mathrm{BMO}_{\mathrm{flag}}$. The proofs of Theorems 30, 32, 33, and 34 will all be given in Section 7.

(5) In Section 8, we prove the Calderón–Zygmund decomposition in the flag two-parameter setting (Theorem 35) and then derive an interpolation result, Theorem 36.

(6) In Section 9, we show that flag singular integrals are not in general bounded from the classical one-parameter Hardy space $H^1(\mathbb{H}^n)$ on the Heisenberg group to $L^1(\mathbb{H}^n)$.

4. L^p estimates for the Littlewood–Paley square function

The purpose of this section is to show that the L^p norm of f is equivalent to the L^p norm of $g_{\mathrm{flag}}(f)$ when $1 < p < \infty$. This was shown in [Müller et al. 1996, Proposition 4.1] for a function $g_{\mathrm{flag}}(f)$ only slightly different than that used here. Our proof is similar in spirit to that work.

Proof of Theorem 6. The proof is similar to that in the pure product case given in [Fefferman and Stein 1982], and follows from iteration and standard vector-valued Littlewood–Paley inequalities. To see this, define

$$L^p(\mathbb{H}^n) \ni f \to F \in H = \ell^2$$

by $F(z, u) = \{\psi_j^{(1)} * f(z, u)\}$, so that

$$\|F\|_H = \left\{ \sum_j |\psi_j^{(1)} * f(z, u)|^2 \right\}^{\frac{1}{2}}.$$

For z fixed, set

$$\tilde{g}(F)(z, u) = \left\{ \sum_k \|\psi_k^{(2)} *_2 F(z, \cdot)(y)\|_H^2 \right\}^{\frac{1}{2}}.$$

It is then easy to see that $\tilde{g}(F)(z, u) = g_{\text{flag}}(f)(z, u)$. For z fixed, by the vector-valued Littlewood–Paley inequality,

$$\int_{\mathbb{H}^n} \tilde{g}(F)^p(z, u)\, dz\, du \le C \int_{\mathbb{H}^n} \|F\|_H^p\, dz\, du.$$

However, $\|F\|_H^p = \{\sum_j : \psi_j^{(1)} * f(z, y)|^2\}^{p/2}$, so integrating with respect to z together with the standard Littlewood–Paley inequality yields

$$\int_{\mathbb{C}^n} \int_{\mathbb{R}} g_{\text{flag}}(f)^p(z, u)\, dz\, du \le C \int_{\mathbb{C}^n} \int_{\mathbb{R}} \left\{ \sum_j |\psi_j^{(1)} * f(z, u)|^2 \right\}^{p/2} dz\, du \le C\|f\|_{L^p(\mathbb{H}^n)}^p,$$

which shows that $\|g_{\text{flag}}(f)\|_p \le C\|f\|_p$.

The proof of the estimate $\|f\|_p \le C\|g_{\text{flag}}(f)\|_p$ is a routine duality argument using the Calderón reproducing formula on $L^2(\mathbb{H}^n)$, for all $f \in L^2 \cap L^p$, $g \in L^2 \cap L^{p'}$ and $1/p + 1/p' = 1$, and the inequality $\|g_{\text{flag}}(f)\|_p \le C\|f\|_p$, which was just proved. This completes the proof of Theorem 6. $\qquad\square$

As in Theorem 6, let $\psi^{(1)} \in \mathscr{S}(\mathbb{H}^n)$ be supported in the unit ball in $\mathbb{H}^n$ and $\psi^{(2)} \in \mathscr{S}(\mathbb{R})$ be supported in the unit ball of $\mathbb{R}$ and satisfy

$$\int_0^\infty |\widehat{\psi^{(2)}}(t\eta)|^4 \frac{dt}{t} = 1$$

for all $\eta \in \mathbb{R} \setminus \{0\}$. We define $\psi^{\natural}(z, u, v) = \psi^{(1)}(z, u)\psi^{(2)}(v)$. Set $\psi_s^{(1)}(z, u) = s^{-n-2}\psi^{(1)}(z/s, u/s^2)$, $\psi_t^{(2)}(v) = t^{-1}\psi(z/t)$ and

$$\psi_{s,t}(z, u) = \int_{\mathbb{R}} \psi_s^{(1)}(z, u-v)\psi_t^{(2)}(v)\, dv.$$

Repeating the proof of Theorem 6, we get, for $1 < p < \infty$,

$$\left\| \left\{ \int_0^\infty \int_0^\infty |\psi_{s,t} * f(z, u)|^2 \frac{dt}{t} \frac{ds}{s} \right\}^{\frac{1}{2}} \right\|_p \le C\|f\|_p$$

and

$$\|f\|_p \approx \left\| \left\{ \int_0^\infty \int_0^\infty |\psi_{s,t} * \psi_{s,t} * f(z, y)|^2 \frac{dt}{t} \frac{ds}{s} \right\}^{\frac{1}{2}} \right\|_p. \tag{4-1}$$

The L^p boundedness of flag singular integrals for $1 < p < \infty$ is then an easy consequence of Theorem 6. This theorem was originally obtained in [Müller et al. 1995] using a different proof that involved the method of transference.

Theorem 38. *Suppose that T is a flag singular integral defined on $\mathbb{H}^n$ with flag kernel $K(z, u)$ as in Definition 7 above. Then T is bounded on L^p for $1 < p < \infty$. Moreover, there exists a constant C depending on p such that, for $f \in L^p$,*

$$\|Tf\|_p \le C\|f\|_p, \quad 1 < p < \infty.$$

Proof. We may first assume that K is an integrable function and then prove the L^p boundedness of T is independent of the L^1 norm of K. The conclusion for general K then follows by an argument used in [Müller et al. 1995]. For all $f \in L^p$, by (4-1),

$$\|T(f)\|_p \leq C \left\| \left\{ \int_0^\infty \int_0^\infty |\psi_{s,t} * \psi_{s,t} * K * f|^2 \frac{dt}{t} \frac{ds}{s} \right\}^{\frac{1}{2}} \right\|_p. \tag{4-2}$$

Now we claim the following estimate: for $f \in L^p$,

$$|\psi_{s,t} * K * f(z,u)| \leq C M_S(f)(z,u), \tag{4-3}$$

where C is a constant which is independent of the L^1 norm of K and $M_S(f)$ is the strong maximal function of f defined in (1-1).

Assuming (4-3) for the moment, we obtain from (4-2) that

$$\|Tf\|_p \leq C \left\| \left\{ \int_0^\infty \int_0^\infty (M_S(\psi_{s,t} * f))^2 \frac{dt}{t} \frac{ds}{s} \right\}^{\frac{1}{2}} \right\|_p \leq C \|f\|_p,$$

where the last inequality follows from the Fefferman–Stein vector-valued maximal inequality.

We now turn to the claim (4-3). This follows from dominating $|\psi_{s,t} * K * f|$ by a product Poisson integral $\mathbb{P}_{\text{prod}} f$, and then dominating the product Poisson integral $\mathbb{P}_{\text{prod}} f$ by the strong maximal function $M_S f$. The arguments are familiar and we leave them to the reader. $\quad\square$

5. Developing the wavelet Calderón reproducing formula

In this section, we develop the wavelet Calderón reproducing formula and the Plancherel–Pólya-type inequalities on test function spaces. These are the main tools used in establishing the theory of Hardy spaces associated with the flag dilation structure. In order to establish the wavelet Calderón reproducing formula and the Plancherel–Pólya-type inequalities, we use the continuous version of the Calderón reproducing formula on $L^2(\mathbb{H}^n)$ and the almost-orthogonality estimates.

We now start the relatively long proof of Theorem 17, beginning with the Calderón reproducing formula in (2-1) that holds for $f \in L^2(\mathbb{H}^n)$ and converges in $L^2(\mathbb{H}^n)$. For any given $\alpha > 0$, we discretize it as

$$
\begin{aligned}
f(z,u) &= \int_0^\infty \int_0^\infty \check{\psi}_{s,t} *_{\mathbb{H}^n} \psi_{s,t} *_{\mathbb{H}^n} f(z,u) \frac{ds}{s} \frac{dt}{t} \\
&= \sum_{j,k \in \mathbb{Z}} \int_{2^{-\alpha(j+1)}}^{2^{-\alpha j}} \int_{2^{-2\alpha(k+1)}}^{2^{-2\alpha k}} \check{\psi}_{s,t} * \psi_{s,t} * f(z,u) \frac{dt}{t} \frac{ds}{s} \\
&= c_\alpha \sum_{j \leq k} \check{\psi}_{j,k} * \psi_{j,k} * f(z,u) + c_\alpha \sum_{j > k} \check{\psi}_{j,k} * \psi_{j,k} * f(z,u) \\
&\quad + \sum_{j,k \in \mathbb{Z}} \int_{2^{-\alpha(j+1)}}^{2^{-\alpha j}} \int_{2^{-2\alpha(k+1)}}^{2^{-2\alpha k}} \{ \check{\psi}_{s,t} * \psi_{s,t} - \check{\psi}_{j,k} * \psi_{j,k} \} * f(z,u) \frac{dt}{t} \frac{ds}{s} \\
&= T_\alpha^{(1)} f(z,u) + T_\alpha^{(2)} f(z,u) + R_\alpha f(z,u),
\end{aligned}
$$

where

$$\psi_{j,k} = \psi_{2^{-\alpha j}, 2^{-2\alpha k}}, \qquad c_\alpha = \int_{2^{-\alpha(j+1)}}^{2^{-\alpha j}} \int_{2^{-2\alpha(k+1)}}^{2^{-2\alpha k}} \frac{dt}{t} \frac{ds}{s} = \ln \frac{2^{-\alpha j}}{2^{-\alpha(j+1)}} \ln \frac{2^{-2\alpha k}}{2^{-2\alpha(k+1)}} = 2(\alpha \ln 2)^2.$$

Notation 39. We have relabeled $\psi_{2^{-\alpha j}, 2^{-2\alpha k}}$ as simply $\psi_{j,k}$ when we replace integrals $\int_0^\infty \int_0^\infty (ds/s)(dt/t)$ by sums $\sum_{j,k \in \mathbb{Z}}$. This abuse of notation should not cause confusion as we will always use j, k, j', k' as subscripts for the discrete components $\psi_{j,k}$, while we always use s, t, s', t' as subscripts for the continuous components $\psi_{s,t}$. Note however that directions are *reversed* in passing from $s, t \in (0, \infty)$ to $j, k \in \mathbb{Z}$, in the sense that $s = 2^{-\alpha j}$ and $t = 2^{-2\alpha k}$ decrease as j and k increase.

To continue we choose a large positive integer N to be fixed later. We decompose the first term $T_\alpha^{(1)} f(z, u)$ by writing the Heisenberg group $\mathbb{H}^n$ as a pairwise disjoint union of dyadic cubes $\mathfrak{Q}$ of side length $2^{-\alpha(j+N)}$, that is,

$$\mathfrak{Q} \in \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(j+N)}).$$

We decompose the second term $T_\alpha^{(2)} f(z, u)$ by writing the Heisenberg group $\mathbb{H}^n$ as a pairwise disjoint union of dyadic rectangles $\mathfrak{R}$ of dimension $2^{-\alpha(j+N)} \times 2^{-2\alpha(k+N)}$, that is, $\mathfrak{R} \in \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(k+N)})$. Recall that

$$\mathsf{R}(j, k) \equiv \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(k+N)}),$$
$$\mathsf{Q}(j) \equiv \mathcal{R}(2^{-\alpha(j+N)} \times 2^{-2\alpha(j+N)}),$$

and that $(z_\mathfrak{Q}, u_\mathfrak{Q})$ is any *fixed* point in the cube $\mathfrak{Q} \in \mathsf{Q}(j)$, and that $(z_\mathfrak{R}, u_\mathfrak{R})$ is any *fixed* point in the rectangle $\mathfrak{R} \in \mathsf{R}(j, k)$.

We further discretize the terms $T_\alpha^{(1)} f(z, u)$ and $T_\alpha^{(2)} f(z, u)$ in different ways, exploiting the one-parameter structure of the Heisenberg group for $T_\alpha^{(1)}$, and exploiting the implicit product structure for $T_\alpha^{(2)}$. We rewrite $T_\alpha^{(1)} f(z, u)$ as

$$\begin{aligned} T_\alpha^{(1)} f(z, u) &= c_\alpha \sum_{j \leq k} \check{\psi}_{j,k} * \psi_{j,k} * f(z, u) \\ &= c_\alpha \sum_{j \leq k} (\check{\psi}_j^{(1)} *_2 \check{\psi}_k^{(2)}) * (\psi_j^{(1)} *_2 \psi_k^{(2)}) * f(z, u) \\ &= c_\alpha \sum_{j \leq k} (\check{\psi}_j^{(1)} *_2 \check{\psi}_k^{(2)} *_2 \psi_k^{(2)}) * \psi_j^{(1)} * f(z, u) \\ &= c_\alpha \sum_{j \in \mathbb{Z}} \left(\check{\psi}_j^{(1)} *_2 \left(\sum_{k \geq j} \check{\psi}_k^{(2)} *_2 \psi_k^{(2)} \right) \right) * \psi_j^{(1)} * f(z, u) \\ &= c_\alpha \sum_{j \in \mathbb{Z}} \check{\psi}_j * \psi_j * f(z, u), \end{aligned}$$

where

$$\psi_j \equiv \psi_j^{(1)} \quad \text{and} \quad \check{\psi}_j \equiv \check{\psi}_j^{(1)} *_2 \left(\sum_{k \geq j} \check{\psi}_k^{(2)} *_2 \psi_k^{(2)} \right). \tag{5-1}$$

Remark 40. It is a standard exercise to prove that $\check{\psi}_j$ satisfies the same type of estimates as does $\psi_j^{(1)}$ on the Heisenberg group $\mathbb{H}^n$.

Now we write

$$T_\alpha^{(1)} f(z, u) = \sum_{j \leq k} \sum_{\mathcal{Q} \in \mathsf{Q}(j)} f_\mathcal{Q} \psi_\mathcal{Q}(z, u) + R_{\alpha, N}^{(1)} f(z, u),$$

$$T_\alpha^{(2)} f(z, u) = \sum_{j > k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} f_\mathcal{R} \psi_\mathcal{R}(z, u) + R_{\alpha, N}^{(2)} f(z, u),$$

where

$$f_\mathcal{Q} \equiv c_\alpha |\mathcal{Q}| \psi_{j,k} * f(z_\mathcal{Q}, u_\mathcal{Q}) \qquad \text{for } \mathcal{Q} \in \mathsf{Q}(j) \text{ and } k \geq j,$$

$$f_\mathcal{R} \equiv c_\alpha |\mathcal{R}| \psi_{j,k} * f(z_\mathcal{R}, u_\mathcal{R}) \qquad \text{for } \mathcal{R} \in \mathsf{R}(j, k) \text{ and } k < j,$$

$$\psi_\mathcal{Q}(z, u) = \frac{1}{|\mathcal{Q}|} \int_\mathcal{Q} \check{\psi}_{j,k}((z, u) \circ (z', u')^{-1}) \, dz' du' \quad \text{for } \mathcal{Q} \in \mathsf{Q}(j) \text{ and } k \geq j,$$

$$\psi_\mathcal{R}(z, u) = \frac{1}{|\mathcal{R}|} \int_\mathcal{R} \check{\psi}_{j,k}((z, u) \circ (z', u')^{-1}) \, dz' du' \quad \text{for } \mathcal{R} \in \mathsf{R}(j, k) \text{ and } k < j,$$

and

$$R_{\alpha, N}^{(1)} f(z, u) = c_\alpha \sum_{j \leq k} \sum_{\mathcal{Q} \in \mathsf{Q}(j)} \int_\mathcal{Q} \check{\psi}_{j,k}((z, u) \circ (z', u')^{-1}) \times [\psi_{j,k} * f(z', u') - \psi_{j,k} * f(z_\mathcal{Q}, u_\mathcal{Q})] \, dz' du',$$

$$R_{\alpha, N}^{(2)} f(z, u) = c_\alpha \sum_{j > k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} \int_\mathcal{R} \check{\psi}_{j,k}((z, u) \circ (z', u')^{-1}) \times [\psi_{j,k} * f(z', u') - \psi_{j,k} * f(z_\mathcal{R}, u_\mathcal{R})] \, dz' du'.$$

Altogether we have

$$f(z, u) = \sum_{j \in \mathbb{Z}} \sum_{\mathcal{Q} \in \mathsf{Q}(j)} f_\mathcal{Q} \psi_\mathcal{Q}(z, u) + \sum_{j > k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} f_\mathcal{R} \psi_\mathcal{R}(z, u)$$
$$+ \{R_\alpha f(z, u) + R_{\alpha, N}^{(1)} f(z, u) + R_{\alpha, N}^{(2)} f(z, u)\}. \quad (5\text{-}2)$$

Recall that we denote by $\mathsf{Q} \equiv \bigcup_{j \in \mathbb{Z}} \mathsf{Q}(j)$ the collection of *all* dyadic cubes, and by $\mathsf{R}_{\text{vert}} \equiv \bigcup_{j > k} \mathsf{R}(j, k)$ the collection of *all strictly vertical* dyadic rectangles. Then we can rewrite (5-2) as

$$f(z, u) = \sum_{\mathcal{Q} \in \mathsf{Q}} f_\mathcal{Q} \psi_\mathcal{Q}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} f_\mathcal{R} \psi_\mathcal{R}(z, u) + \{R_\alpha + R_{\alpha, N}^{(1)} + R_{\alpha, N}^{(2)}\} f(z, u), \quad (5\text{-}3)$$

which is a precursor to the *wavelet* form of the Calderón reproducing formula given in the statement of Theorem 17.

The following theorem is the analogue of [Han 1994, Theorem 1.19] for the operators R_α, $R_{\alpha, N}^{(1)}$, and $R_{\alpha, N}^{(2)}$.

Theorem 41. *For fixed M and $0 < \delta < 1$, we can choose M′ and $0 < \alpha < \varepsilon$ sufficiently small, and then choose N sufficiently large, so that the operators R_α, $R_\alpha^{(1)}$, and $R_\alpha^{(2)}$ satisfy*

$$\|R_\alpha f\|_{L^p(\mathbb{H}^n)} + \|R_{\alpha, N}^{(1)} f\|_{L^p(\mathbb{H}^n)} + \|R_{\alpha, N}^{(2)} f\|_{L^p(\mathbb{H}^n)} \leq \tfrac{1}{2} \|f\|_{L^p(\mathbb{H}^n)}, \quad f \in L^p(\mathbb{H}^n), \ 1 < p < \infty,$$

$$\|R_\alpha f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)} + \|R_{\alpha, N}^{(1)} f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)} + \|R_{\alpha, N}^{(2)} f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)} \leq \tfrac{1}{2} \|f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)}, \quad f \in \mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n). \tag{5-4}$$

With Theorem 41 in hand, we obtain that the operator

$$S_{\alpha,N} \equiv I - R_\alpha - R_{\alpha,N}^{(1)} f - R_{\alpha,N}^{(2)}$$

is bounded and invertible on $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$. It follows that, with $\tilde{\psi}_{\mathcal{Q}} \equiv S_{\alpha,N}^{-1} \psi_{\mathcal{Q}}$ and $\tilde{\psi}_{\mathcal{R}} \equiv S_{\alpha,N}^{-1} \psi_{\mathcal{R}}$,

$$f(z, u) = \sum_{\mathcal{Q} \in \mathsf{Q}} f_{\mathcal{Q}} \tilde{\psi}_{\mathcal{Q}}(z, u) + \sum_{\mathcal{R} \in \mathsf{R}_{\text{vert}}} f_{\mathcal{R}} \tilde{\psi}_{\mathcal{R}}(z, u), \quad f \in \mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n), \tag{5-5}$$

where $\tilde{\psi}_{\mathcal{Q}}$ and $\tilde{\psi}_{\mathcal{R}}$ are in $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$, and the convergence in (5-5) is in both $L^p(\mathbb{H}^n)$ and in the Banach space $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$. This finally is the *wavelet* form of the Calderón reproducing formula given in the statement of Theorem 17. The same argument shows that (5-5) holds for $f \in L^p(\mathbb{H}^n)$ with convergence in $L^p(\mathbb{H}^n)$, provided $1 < p < \infty$. In fact we obtain that (5-5) holds for f in any Banach space $\mathcal{X}(\mathbb{H}^n)$ with convergence in $\mathcal{X}(\mathbb{H}^n)$, provided we have operator bounds

$$\|R_\alpha f\|_{\mathcal{X}(\mathbb{H}^n)} + \|R_{\alpha,N}^{(1)} f\|_{\mathcal{X}(\mathbb{H}^n)} + \|R_{\alpha,N}^{(2)} f\|_{\mathcal{X}(\mathbb{H}^n)} \leq \tfrac{1}{2} \|f\|_{\mathcal{X}(\mathbb{H}^n)}, \quad f \in \mathcal{X}(\mathbb{H}^n).$$

We turn first to proving the molecular estimates in (5-4), but only for

$$\|R_{\alpha,N}^{(1)} f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)} \quad \text{and} \quad \|R_{\alpha,N}^{(2)} f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)},$$

as the estimate for $\|R_\alpha f\|_{\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)}$ is similar, but easier. We will use the following special $T1$-type theorem on the Heisenberg group $\mathbb{H}^n$ (see [Han 1998; 1994] for the Euclidean case) to prove a corresponding product version below. Recall the definition of the one-parameter molecular space $\mathcal{M}^{M'+\delta}(\mathbb{H}^n)$.

Definition 42. Let $M' \in \mathbb{N}$ be a positive integer, $0 < \delta \leq 1$, and let $Q = 2n + 2$ denote the homogeneous dimension of $\mathbb{H}^n$. The *one-parameter* molecular space $\mathcal{M}^{M'+\delta}(\mathbb{H}^n)$ consists of all functions $f(z, u)$ on $\mathbb{H}^n$ satisfying the moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta f(z, u) \, dz \, du = 0 \quad \text{for all } |\alpha| + 2|\beta| \leq M',$$

and such that there is a nonnegative constant A satisfying the differential inequalities

$$|\partial_z^\alpha \partial_u^\beta f(z, u)| \leq A \frac{1}{(1 + |z|^2 + |u|)^{(Q+M'+|\alpha|+2|\beta|+\delta)/2}} \quad \text{for all } |\alpha| + 2|\beta| \leq M'$$

and

$$|\partial_z^\alpha \partial_u^\beta f(z, u) - \partial_z^\alpha \partial_u^\beta f(z', u')| \leq A \frac{|(z, u) \circ (z', u')^{-1}|^\delta}{(1 + |z|^2 + |u|)^{(Q+M+\delta+M'+2\delta)/2}}$$

$$\text{for all } |\alpha| + 2|\beta| = M' \text{ and } |(z, u) \circ (z', u')^{-1}| \leq \tfrac{1}{2}(1 + |z|^2 + |u|)^{\frac{1}{2}}.$$

Theorem 43. *Suppose* $T : L^2(\mathbb{H}^n) \to L^2(\mathbb{H}^n)$ *is a bounded linear operator with kernel* $K((z, u), (z', u'))$; *that is,*

$$Tf(z, u) = \int_{\mathbb{H}^n} K((z, u), (z', u')) f(z', u') \, dz' \, du'.$$

Suppose furthermore that K satisfies

$$\int_{\mathbb{H}^n} z^\alpha u^\beta K((z, u), (z', u')) \, dz \, du = 0,$$

$$\int_{\mathbb{H}^n} (z')^\alpha (u')^\beta K((z, u), (z', u')) \, dz' \, du' = 0$$

for all $0 \le |\alpha|, \beta$, and

$$|\partial_z^\alpha \partial_u^\beta \partial_{z'}^{\alpha'} \partial_{u'}^{\beta'} K((z, u), (z', u'))| \le A \frac{1}{|(z, u) \circ (z', u')^{-1}|^{Q + |\alpha| + 2\beta + |\alpha'| + 2\beta'}}$$

for all $0 \le |\alpha|, \beta, |\alpha'|, \beta'$. Then

$$T : L^p(\mathbb{H}^n) \to L^p(\mathbb{H}^n) \qquad \text{for } 1 < p < \infty,$$

$$T : \mathcal{M}^{M'+\delta}(\mathbb{H}^n) \to \mathcal{M}^{M'+\delta}(\mathbb{H}^n) \quad \text{for all } M' \text{ and } 0 < \delta < 1,$$

and, moreover, the operator norms satisfy

$$\|T\|_{L^p(\mathbb{H}^n)} \le C_p A \quad \text{and} \quad \|T\|_{\mathcal{M}^{M'+\delta}(\mathbb{H}^n)} \le C_{M', \delta} A.$$

We will use the technique of lifting to the product space $\mathcal{M}^{M'+\delta}_{\text{product}}(\mathbb{H}^n \times \mathbb{R})$ together with the following special product $T1$-type theorem on the product group $\mathbb{H}^n \times \mathbb{R}$.

Theorem 44. *Suppose that $T : L^2(\mathbb{H}^n \times \mathbb{R}) \to L^2(\mathbb{H}^n \times \mathbb{R})$ is a bounded linear operator with kernel $K([(z, u), v], [(z', u'), v'])$; that is,*

$$Tf((z, u), v) = \int_{\mathbb{H}^n \times \mathbb{R}} K([(z, u), v], [(z', u'), v']) f((z', u'), v') \, dz' \, du' \, dv'.$$

Suppose furthermore that K satisfies

$$\int_{\mathbb{H}^n} z^\alpha u^\beta K([(z, u), v], [(z', u'), v']) \, dz \, du = 0,$$

$$\int_{\mathbb{H}^n} (z')^\alpha (u')^\beta K([(z, u), v], [(z', u'), v']) \, dz' \, du' = 0,$$

$$\int_{\mathbb{R}} v^\gamma K([(z, u), v], [(z', u'), v']) \, dv = 0,$$

$$\int_{\mathbb{R}} (v')^\gamma K([(z, u), v], [(z', u'), v']) \, dv' = 0$$

for all $0 \le |\alpha|, \beta, \gamma$, and

$$|\partial_z^\alpha \partial_u^\beta \partial_v^\gamma \partial_{z'}^{\alpha'} \partial_{u'}^{\beta'} \partial_{v'}^{\gamma'} K([(z, u), v], [(z', u'), v'])| \le A \frac{1}{|(z, u) \circ (z', u')^{-1}|^{Q + |\alpha| + 2\beta + |\alpha'| + 2\beta'}} \frac{1}{|v - v'|^{1 + \gamma_1 + \gamma_2}}$$

for all $0 \le |\alpha|, \beta, \gamma, |\alpha'|, \beta', \gamma'$. Then

$$T : L^p(\mathbb{H}^n \times \mathbb{R}) \to L^p(\mathbb{H}^n \times \mathbb{R}) \qquad \text{for } 1 < p < \infty,$$

$$T : \mathcal{M}^{M'+\delta}_{\text{product}}(\mathbb{H}^n \times \mathbb{R}) \to \mathcal{M}^{M'+\delta}_{\text{product}}(\mathbb{H}^n \times \mathbb{R}) \quad \text{for all } M' \text{ and } 0 < \delta < 1,$$

and, moreover, the operator norms satisfy

$$\|T\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \le C_p A \quad and \quad \|T\|_{\mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})} \le C_{M',\alpha} A.$$

We postpone the proofs of these $T1$-type theorems, and turn now to using them to complete the proof of Theorem 41, which in turn completes the proof of Theorem 17.

5.1. *Boundedness on the flag molecular space.* We prove the estimates for the operators $R_{\alpha,N}^{(1)}$ and $R_{\alpha,N}^{(2)}$ in Theorem 41 separately, beginning with $R_{\alpha,N}^{(2)}$.

5.1.1. *The operator $R_{\alpha,N}^{(2)}$.* Here we prove the boundedness of the error operator

$$R_{\alpha,N}^{(2)} f(z,u) = c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} \int_{\mathcal{R}} \check{\psi}_{j,k}((z,u) \circ (z',u')^{-1}) \times [\psi_{j,k} * f(z',u') - \psi_{j,k} * f(z_{\mathcal{R}}, u_{\mathcal{R}})] \, dz' \, du'$$

on the flag molecular space $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$, where M' is taken sufficiently small compared to M as in the component functions. We begin by lifting the desired inequality to the product group $\mathbb{H}^n \times \mathbb{R}$ and reducing matters to Theorem 44. So we begin by writing

$$
\begin{aligned}
R_{\alpha,N}^{(2)} & f(z,u) \\
&= c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} \int_{\mathcal{R}} \check{\psi}_{j,k}((z,u) \circ (z',u')^{-1}) \\
&\quad \times \int [\psi_{j,k}((z',u') \circ (z'',u'')^{-1}) - \psi_{j,k} * f((z_{\mathcal{R}}, u_{\mathcal{R}}) \circ (z'',u'')^{-1})] f(z'',u'') \, dz'' \, du'' \, dz' \, du' \\
&= c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} \int_{\mathcal{R}} \left\{ \int \check{\psi}_j^{(1)}(z - z', u - u' + \operatorname{Im} z\bar{z}' - w) \check{\psi}_k^{(2)}(w) \, dw \right\} \\
&\quad \times \int \left\{ \int \psi_j^{(1)}(z' - z'', u' - u'' + \operatorname{Im} z'\bar{z}'' - w') \check{\psi}_k^{(2)}(w') \right. \\
&\quad \left. - \int \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}}\bar{z}'' - w') \psi_k^{(2)}(w') \right\} dw' \int F(z'', u'' - w'', w'') \, dw'',
\end{aligned}
$$

where

$$f(z,u) = \pi F(z,u) = \int F((z,u-w),w) \, dw$$

and $F((z,u),w) \in \mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})$. We continue with

$$
\begin{aligned}
R_{\alpha,N}^{(2)} f(z,u) = c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in \mathsf{R}(j,k)} \int_{\mathcal{R}} \iiiint \check{\psi}_j^{(1)}(z - z', u - u' + \operatorname{Im} z\bar{z}' - w) \check{\psi}_k^{(2)}(w) \\
\times \{\psi_j^{(1)}(z' - z'', u' - u'' + \operatorname{Im} z'\bar{z}'' - w') - \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}}\bar{z}'' - w')\} \\
\times \check{\psi}_k^{(2)}(w') F(z'', u'' - w'', w'') \, dz'' \, du'' \, dw'' \, dw' \, dw \, dz' \, du'.
\end{aligned}
$$

Now for fixed w'' make the change of variable $u'' \to u'' + w''$ (in the sense that $u'' \to \tilde{u}'' + w''$ and we then rewrite $\tilde{u}''$ as u'') to obtain

$$
\begin{aligned}
R_{\alpha,N}^{(2)} f(z,u) = c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in R(j,k)} \int_{\mathcal{R}} \iiiint \check{\psi}_j^{(1)}(z-z', u-u'+\operatorname{Im} z\overline{z}' - w) \check{\psi}_k^{(2)}(w) \\
\times \{\psi_j^{(1)}(z'-z'', u'-u''-w'+\operatorname{Im} z'\overline{z}'' - w') - \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}} \overline{z}'' - w' - w'')\} \\
\times \check{\psi}_k^{(2)}(w') F(z'', u'', w'') \, dz'' \, du'' \, dw'' \, dw' \, dw \, dz' \, du'.
\end{aligned}
$$

Then, making a change of variable $w' \to w' - w''$ (in the sense of the previous change of variable), we get

$$
\begin{aligned}
R_{\alpha,N}^{(2)} f(z,u) = c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in R(j,k)} \int_{\mathcal{R}} \iiiint \check{\psi}_j^{(1)}(z-z', u-u'+\operatorname{Im} z\overline{z}' - w) \check{\psi}_k^{(2)}(w) \\
\times \{\psi_j^{(1)}(z'-z'', u'-u''+\operatorname{Im} z'\overline{z}'' - w') - \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}} \overline{z}'' - w')\} \\
\times \check{\psi}_k^{(2)}(w' - w'') F(z'', u'', w'') \, dz'' \, du'' \, dw'' \, dw' \, dw \, dz' \, du'.
\end{aligned}
$$

Finally, making the change of variable $w \to w - w'$, we get

$$
\begin{aligned}
R_{\alpha,N}^{(2)} f(z,u) & \\
= c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in R(j,k)} \int_{\mathcal{R}} \iiiint \check{\psi}_j^{(1)}(z-z', u-u'+\operatorname{Im} z\overline{z}' - w + w') \check{\psi}_k^{(2)}(w - w') \\
\times \{\psi_j^{(1)}(z'-z'', u'-u''+\operatorname{Im} z'\overline{z}'' - w') - \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}} \overline{z}'' - w')\} \check{\psi}_k^{(2)}(w' - w'') \\
\times F(z'', u'', w'') \, dz'' \, du'' \, dw'' \, dw' \, dw \, dz' \, du' \\
= \int \widetilde{R}_{\alpha,N}^{(2)} F((z, u-w), w) \, dw,
\end{aligned}
$$

where the kernel of $\widetilde{R}_{\alpha,N}^{(2)}$ is given by

$$
\widetilde{R}_{\alpha,N}^{(2)}[((z,u), w), ((z'', u''), w'')] = c_\alpha \sum_{j>k} \sum_{\mathcal{R} \in R(j,k)} \int_{\mathcal{R}} \int \check{\psi}_j^{(1)}(z-z', u-u'+\operatorname{Im} z\overline{z}' + w') \check{\psi}_k^{(2)}(w - w')
$$

$$
\times \{\psi_j^{(1)}(z'-z'', u'-u''+\operatorname{Im} z'\overline{z}'' - w') - \psi_j^{(1)}(z_{\mathcal{R}} - z'', u_{\mathcal{R}} - u'' + \operatorname{Im} z_{\mathcal{R}} \overline{z}'' - w')\} \check{\psi}_k^{(2)}(w' - w'') \, dz' \, du' \, dw'.
$$

Now it suffices to show that

$$
\widetilde{R}_{\alpha,N}^{(2)} F \in \mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})
$$

with small norm, since we then conclude that

$$
R_{\alpha,N}^{(2)} f \in \mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)
$$

with small norm. To do this we need only check that the kernel of $\widetilde{R}_{\alpha,N}^{(2)}$ satisfies the conditions of Theorem 44 with small bounds.

For this we rewrite the kernel in terms of Heisenberg group multiplication as

$$
\widetilde{R}_{\alpha,N}^{(2)}[((z,u),w),((z'',u''),w'')] = c_\alpha \sum_{j>k} \sum_{\mathscr{R}\in\mathsf{R}(j,k)} \int_{\mathscr{R}} \int \check{\psi}_j^{(1)}((z,u)\circ(z',u'-w')^{-1})\check{\psi}_k^{(2)}(w-w')
$$

$$
\times \{\psi_j^{(1)}((z',u'-w')\circ(z'',u'')^{-1}) - \psi_j^{(1)}((z_{\mathscr{R}},u_{\mathscr{R}}-w')\circ(z'',u'')^{-1})\}\psi_k^{(2)}(w'-w'')\,dz'\,du'\,dw'.
$$

By construction we have

$$
\psi_j^{(1)}((z',u'-w')\circ(z'',u'')^{-1}) - \psi_j^{(1)}((z_{\mathscr{R}},u_{\mathscr{R}}-w')\circ(z'',u'')^{-1}) \sim 2^{-N}\psi_j^{(1)}((z',u'-w')\circ(z'',u'')^{-1}),
$$

in the sense that the left side satisfies the same moment, size and smoothness conditions as the right side. Thus we have

$$
\sum_{\mathscr{R}\in\mathsf{R}(j,k)} \int_{\mathscr{R}} \int \check{\psi}_j^{(1)}((z,u)\circ(z',u'-w')^{-1})
$$

$$
\times \{\psi_j^{(1)}((z',u'-w')\circ(z'',u'')^{-1}) - \psi_j^{(1)}((z_{\mathscr{R}},u_{\mathscr{R}}-w')\circ(z'',u'')^{-1})\}\,dz'\,du'
$$

$$
\sim \sum_{\mathscr{R}\in\mathsf{R}(j,k)} \int_{\mathscr{R}} \int \check{\psi}_j^{(1)}((z,u)\circ(z',u'-w')^{-1})2^{-N}\psi_j^{(1)}((z',u'-w')\circ(z'',u'')^{-1})\,dz'\,du'
$$

$$
\sim 2^{-N}\psi_j^{(1)}((z,u)\circ(z'',u'')^{-1}). \tag{5-6}
$$

We also have

$$
\int \check{\psi}_k^{(2)}(w-w')\psi_k^{(2)}(w'-w'')\,dw' \sim \psi_k^{(2)}(w-w'').
$$

So altogether we obtain

$$
\widetilde{R}_{\alpha,N}^{(2)}[((z,u),w),((z'',u''),w'')] \sim 2^{-N}\sum_{j>k}\psi_j^{(1)}((z,u)\circ(z'',u'')^{-1})\psi_k^{(2)}(w-w''),
$$

which satisfies the hypotheses of Theorem 44 with bounds roughly 2^{-N}, since $\psi^{(1)} \in \mathscr{S}(\mathbb{H}^n)$ and $\psi^{(2)} \in \mathscr{S}(\mathbb{R})$. Here we are using the well-known fact that the partial sums $\sum_{j<M}\psi_j$ of an approximate identity satisfy Calderón–Zygmund kernel conditions of infinite order *uniformly* in M.

5.1.2. *The operator* $R_{\alpha,N}^{(1)}$. Now we turn to boundedness of the error operator

$$
R_{\alpha,N}^{(1)}f(z,u) = c_\alpha \sum_{j\le k} \sum_{\mathscr{Q}\in\mathsf{Q}(j)} \int_{\mathscr{Q}} \check{\psi}_{j,k}((z,u)\circ(z',u')^{-1})[\psi_{j,k}*f(z',u') - \psi_{j,k}*f(z_{\mathscr{Q}},u_{\mathscr{Q}})]\,dz'\,du',
$$

on the flag molecular space $\mathcal{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$, where M' is taken sufficiently small compared to M as in the component functions. Applying the calculation used for the term $R_{\alpha,N}^{(2)}$ above, we can obtain

$$
R_{\alpha,N}^{(1)}f(z,u) = \int \widetilde{R}_{\alpha,N}^{(1)}F((z,u-w),w)\,dw,
$$

where the kernel of $\widetilde{R}_{\alpha,N}^{(1)}$ is given by

$$\widetilde{R}_{\alpha,N}^{(1)}[((z,u),w),((z'',u''),w'')] = c_\alpha \sum_{j \leq k} \sum_{\mathfrak{Q} \in \mathsf{Q}(j)} \int_{\mathfrak{Q}} \int \check{\psi}_j^{(1)}(z-z',u-u'+\mathrm{Im}\, z\bar{z}'+w')\check{\psi}_k^{(2)}(w-w')$$

$$\times \{\psi_j^{(1)}(z'-z'',u'-u''+\mathrm{Im}\, z'\bar{z}''-w')-\psi_j^{(1)}(z_{\mathfrak{R}}-z'',u_{\mathfrak{R}}-u''+\mathrm{Im}\, z_{\mathfrak{R}}\bar{z}''-w')\}\check{\psi}_k^{(2)}(w'-w'')\,dz'\,du'\,dw'.$$

By construction we have

$$\psi_j^{(1)}((z',u'-w') \circ (z'',u'')^{-1})-\psi_j^{(1)}((z_{\mathfrak{R}},u_{\mathfrak{R}}-w') \circ (z'',u'')^{-1}) \sim 2^{-N} \psi_j^{(1)}((z',u'-w') \circ (z'',u'')^{-1}),$$

in the sense that the left side satisfies the same moment, size, and smoothness conditions as the right side. Thus we have

$$\sum_{\mathfrak{Q} \in \mathsf{Q}(j)} \int_{\mathfrak{Q}} \int \check{\psi}_j^{(1)}((z,u) \circ (z',u'-w')^{-1})$$

$$\times \{\psi_j^{(1)}((z',u'-w') \circ (z'',u'')^{-1})-\psi_j^{(1)}((z_{\mathfrak{R}},u_{\mathfrak{R}}-w') \circ (z'',u'')^{-1})\}\,dz'\,du'$$

$$\sim \sum_{\mathfrak{Q} \in \mathsf{Q}(j)} \int_{\mathfrak{R}} \int \check{\psi}_j^{(1)}((z,u) \circ (z',u'-w')^{-1})2^{-N}\psi_j^{(1)}((z',u'-w') \circ (z'',u'')^{-1})\,dz'\,du'$$

$$\sim 2^{-N}\psi_j^{(1)}((z,u) \circ (z'',u'')^{-1}).$$

We also have

$$\int \check{\psi}_k^{(2)}(w-w')\psi_k^{(2)}(w'-w'')\,dw' \sim \psi_k^{(2)}(w-w'').$$

So altogether we obtain

$$\widetilde{R}_{\alpha,N}^{(1)}[((z,u),w),((z'',u''),w'')] \sim 2^{-N} \sum_{j \leq k} \psi_j^{(1)}((z,u) \circ (z'',u'')^{-1})\psi_k^{(2)}(w-w''),$$

which satisfies the hypotheses of Theorem 44 with bounds roughly 2^{-N}, since $\psi^{(1)} \in \mathscr{S}(\mathbb{H}^n)$ and $\psi^{(2)} \in \mathscr{S}(\mathbb{R})$.

It now follows that the kernels of both $\widetilde{R}_{\alpha,N}^{(1)}$ and $\widetilde{R}_{\alpha,N}^{(2)}$ satisfy the hypotheses of Theorem 44 with bounds roughly 2^{-N}, and we conclude that

$$\|\widetilde{R}_{\alpha,N}^{(i)}F\|_{\mathcal{M}_{\mathrm{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})} \lesssim 2^{-N}\|F\|_{\mathcal{M}_{\mathrm{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})}, \quad i = 1, 2.$$

Thus we obtain, for each $i = 1, 2$,

$$\|R_{\alpha,N}^{(i)}f\|_{\mathcal{M}_{\mathrm{flag}}^{M'+\delta}(\mathbb{H}^n)} \leq \inf_{f = \pi F} \|\widetilde{R}_{\alpha,N}^{(i)}F\|_{\mathcal{M}_{\mathrm{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})} \lesssim 2^{-N} \inf_{f = \pi F} \|F\|_{\mathcal{M}_{\mathrm{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})} = 2^{-N}\|f\|_{\mathcal{M}_{\mathrm{flag}}^{M'+\delta}(\mathbb{H}^n)},$$

and taking N sufficiently large completes the proof of the molecular estimates in (5-4).

5.1.3. *The L^p estimates.* Finally, we turn to proving the L^p estimates in (5-4) for $1 < p < \infty$,

$$\|R_\alpha f\|_{L^p(\mathbb{H}^n)} + \|R_{\alpha,N}^{(1)}f\|_{L^p(\mathbb{H}^n)} + \|R_{\alpha,N}^{(2)}f\|_{L^p(\mathbb{H}^n)} \leq \tfrac{1}{2}\|f\|_{L^p(\mathbb{H}^n)}.$$

The estimates for $R_{\alpha,N}^{(1)}$ and $R_{\alpha,N}^{(2)}$ follow from the estimates established above for the kernels of the lifted operators $\widetilde{R}_{\alpha,N}^{(1)}$ and $\widetilde{R}_{\alpha,N}^{(2)}$. Indeed, for $f \in L^p(\mathbb{H}^n)$, we can use a result in [Müller et al. 1995] to find $F \in L^p(\mathbb{H}^n \times \mathbb{R})$ with $f = \pi F$ and $\|F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \leq C \|f\|_{L^p(\mathbb{H}^n)}$. Then we have

$$\|R_{\alpha,N}^{(i)} f\|_{L^p(\mathbb{H}^n)} \leq \|\widetilde{R}_{\alpha,N}^{(i)} F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \lesssim 2^{-N} \|F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \leq C 2^{-N} \|f\|_{L^p(\mathbb{H}^n)}.$$

In similar fashion, the kernel of the lifted operator $\widetilde{R}_\alpha$ can be shown to satisfy product kernel estimates with constant A that is a multiple of $1 - 2^{-\alpha}$, and so we obtain from Theorem 44 that

$$\|\widetilde{R}_\alpha F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \lesssim (1 - 2^{-\alpha}) \|F\|_{L^p(\mathbb{H}^n \times \mathbb{R})},$$

and hence, with $f = \pi F$ as above,

$$\|R_\alpha f\|_{L^p(\mathbb{H}^n)} \leq \|\widetilde{R}_\alpha F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \lesssim (1 - 2^{-\alpha}) \|F\|_{L^p(\mathbb{H}^n \times \mathbb{R})} \leq C(1 - 2^{-\alpha}) \|f\|_{L^p(\mathbb{H}^n)}.$$

If we now take $0 < \alpha < 1$ sufficiently small, and then N sufficiently large, we obtain the L^p estimates in (5-4). This concludes our proof of Theorem 41.

5.2. *The $T1$-type theorems.* The proof of Theorem 43 in the one-parameter case follows the argument in [Gilbert et al. 2002], where the same result is proved in the Euclidean setting. For this we will need an extension to the Heisenberg group of the generalization of Meyer's lemma by Torres [1991].

Lemma 45. *Suppose $T : L^2(\mathbb{H}^n) \to L^2(\mathbb{H}^n)$ is a bounded linear operator with kernel $K((z, u), (z', u'))$ satisfying the kernel conditions in the hypotheses of* Theorem 43. *Suppose that $M \geq 0$ and that $T((z, u)^{(\alpha'', \beta'')}) = 0$ for all multi-indices (α'', β'') with $|\alpha''| + 2\beta'' \leq M$. Then, for any two points $(z, u), (z'', u'') \in \mathbb{H}^n$ and any smooth φ on $\mathbb{H}^n$ with compact support, and any multi-index (α', β') with $|\alpha'| + 2\beta' = M' \leq M$, we have the identity*

$$\partial_z^{\alpha'} \partial_u^{\beta'} T\varphi(z, u) - \partial_z^{\alpha'} \partial_u^{\beta'} T\varphi(z'', u'')$$
$$= \int \partial_z^{\alpha'} \partial_u^{\beta'} K((z, u), (z', u'))$$
$$\times \left\{ \varphi(z', u') - \sum_{|\alpha''| + 2\beta'' \leq M'} c_{\alpha'', \beta''} \partial_z^{\alpha''} \partial_u^{\beta''} \varphi(z, u) [(z', u') \circ (z, u)^{-1}]^{(\alpha'', \beta'')} \right\} \tilde{\theta}(z', u')\, dz'\, du'$$
$$- \int \partial_z^{\alpha'} \partial_u^{\beta'} K((z'', u''), (z', u'))$$
$$\times \left\{ \varphi(z', u') - \sum_{|\alpha''| + 2\beta'' \leq M'} c_{\alpha'', \beta''} \partial_z^{\alpha''} \partial_u^{\beta''} \varphi(z'', u'') [(z', u') \circ (z'', u'')^{-1}]^{(\alpha'', \beta'')} \right\} \tilde{\theta}(z', u')\, dz'\, du'$$
$$+ \int \{\partial_z^{\alpha'} \partial_u^{\beta'} K((z, u), (z', u')) - \partial_z^{\alpha'} \partial_u^{\beta'} K((z'', u''), (z', u'))\}$$
$$\times \left\{ \varphi(z', u') - \sum_{|\alpha''| + 2\beta'' \leq M'} c_{\alpha'', \beta''} \partial_z^{\alpha''} \partial_u^{\beta''} \varphi(z'', u'') [(z', u') \circ (z'', u'')^{-1}]^{(\alpha'', \beta'')} \right\} (1 - \tilde{\theta}(z', u'))\, dz'\, du'$$
$$+ \sum_{|\alpha''| + 2\beta'' \leq M'} \left\{ c_{\alpha'', \beta''} \partial_z^{\alpha''} \partial_u^{\beta''} \varphi(z, u) - \sum_{|\alpha'''| + 2\beta''' \leq M' - |\alpha''| - 2\beta''} c_{\alpha''', \beta''} \partial_z^{\alpha''' + \alpha''} \partial_u^{2\beta''' + 2\beta''} \right.$$
$$\left. \times \varphi(z'', u'') [(z, u) \circ (z'', u'')^{-1}]^{(\alpha''', \beta''')} \right\} T_{(\alpha'', \beta''), (\alpha', \beta')} \tilde{\theta}(z, u).$$

The proof of this lemma follows verbatim that of [Torres 1991, Lemma 3.1.22, page 62].

With this result in hand, the proof of Theorem 43 follows closely the argument in the Euclidean case in [Gilbert et al. 2002], and the reader can find complete details in [Han et al. 2012].

Proof of Theorem 44. To prove the product version we note that the above one-parameter proof extends virtually verbatim to establish a *vector-valued* version in a Banach space. Indeed, all the main tools, such as integration, differentiation, and Taylor's formula, carry over to the Banach space setting. First we will define the X-valued molecular space $\mathcal{M}^{M+\delta, M_1, M_2}(\mathbb{H}^n; X)$, and then we will give the extension of Theorem 43 to this space.

Definition 46. Let X be a Banach space with norm $|x|$ for $x \in X$. Let $M, M_1, M_2 \in \mathbb{N}$ be positive integers, $0 < \delta \leq 1$, and let $Q = 2n + 2$ denote the homogeneous dimension of $\mathbb{H}^n$. The *one-parameter* molecular space $\mathcal{M}^{M+\delta, M_1, M_2}(\mathbb{H}^n; X)$ consists of all X-valued functions $f : \mathbb{H}^n \to X$ satisfying the moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta f(z, u) \, dz \, du = 0 \quad \text{for all } |\alpha| + 2|\beta| \leq M_1,$$

and such that there is a nonnegative constant A satisfying the differential inequalities

$$|\partial_z^\alpha \partial_u^\beta f(z, u)|_X \leq A \frac{1}{(1 + |z|^2 + |u|)^{(Q + M + |\alpha| + 2|\beta| + \delta)/2}} \quad \text{for all } |\alpha| + 2|\beta| \leq M_2$$

and

$$|\partial_z^\alpha \partial_u^\beta f(z, u) - \partial_z^\alpha \partial_u^\beta f(z', u')|_X \leq A \frac{|(z, u) \circ (z', u')^{-1}|^\delta}{(1 + |z|^2 + |u|)^{(Q + M + \delta + M_2 + 2\delta)/2}}$$

$$\text{for all } |\alpha| + 2|\beta| = M_2 \text{ and } |(z, u) \circ (z', u')^{-1}| \leq \tfrac{1}{2}(1 + |z|^2 + |u|)^{\frac{1}{2}}.$$

We have the following extension of Theorem 43 to X-valued functions for an arbitrary Banach space X.

Theorem 47. *Suppose $T : L^2(\mathbb{H}^n) \to L^2(\mathbb{H}^n)$ is a bounded linear operator with kernel $K((z, u), (z', u'))$; that is,*

$$Tf(z, u) = \int_{\mathbb{H}^n} K((z, u), (z', u')) f(z', u') \, dz' \, du', \quad f \in L^2(\mathbb{H}^n).$$

Suppose furthermore that K satisfies

$$\int_{\mathbb{H}^n} z^\alpha u^\beta K((z, u), (z', u')) \, dz \, du = 0,$$

$$\int_{\mathbb{H}^n} (z')^\alpha (u')^\beta K((z, u), (z', u')) \, dz' \, du' = 0$$

for all $0 \leq |\alpha|, \beta$, and

$$|\partial_z^\alpha \partial_u^\beta \partial_{z'}^{\alpha'} \partial_{u'}^{\beta'} K((z, u), (z', u'))| \lesssim \frac{1}{|(z, u) \circ (z', u')^{-1}|^{Q + |\alpha| + 2\beta + |\alpha'| + 2\beta'}}$$

for all $0 \leq |\alpha|, \beta, |\alpha'|, \beta'$. For $f : \mathbb{H}^n \to X$, we define Tf by the Banach-space-valued integrals

$$Tf(z, u) = \int_{\mathbb{H}^n} K((z, u), (z', u')) f(z', u') \, dz' \, du'.$$

Then

$$T : \mathcal{M}^{M'+\delta}(\mathbb{H}^n; X) \to \mathcal{M}^{M'+\delta}(\mathbb{H}^n; X)$$

is bounded for all M' and $0 < \delta < 1$. Moreover, the operator norm satisfies

$$\|T\|_{\mathcal{M}^{M'+\delta}(\mathbb{H}^n; X)} \le C_{M',\delta}.$$

Proof. We simply repeat the scalar proof of Theorem 43 but replace $|\partial_z^\alpha \partial_u^\beta f(z, u)|$ by $|\partial_z^\alpha \partial_u^\beta f(z, u)|_X$ throughout and use Banach space analogues of Taylor's theorem and the identities of [Torres 1991]. $\square$

Now we can quickly finish the proof of Theorem 44. We take $X = \mathcal{M}^{M'+\delta}(\mathbb{R})$ and note that the identification of product and iterated molecular spaces, namely,

$$\mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R}) = \mathcal{M}^{M'+\delta}\big(\mathbb{H}^n; \mathcal{M}^{M'+\delta}(\mathbb{R})\big) = \mathcal{M}^{M'+\delta}(\mathbb{H}^n; X), \tag{5-7}$$

follows immediately from the definitions of the spaces involved; see Definitions 42 and 10 and the definition of $\mathcal{M}^{M+\delta, M_1, M_2}(\mathbb{R})$, which we recall here.

Definition 48. Let $M \in \mathbb{N}$ be a positive integer and $0 < \delta \le 1$. The *one-parameter* molecular space $\mathcal{M}^{M+\delta, M_1, M_2}(\mathbb{R})$ consists of all functions $f(v)$ on $\mathbb{R}$ satisfying the moment conditions

$$\int_{\mathbb{R}} v^\gamma f(v)\, dv = 0 \quad \text{for all } 2\gamma \le M_1,$$

and such that there is a nonnegative constant A satisfying the differential inequalities

$$|\partial_v^\gamma f(v)| \le A \frac{1}{(1+|v|)^{1+M+\gamma+\delta}} \qquad \text{for all } 2\gamma \le M_2,$$

$$|\partial_v^{M_2} f(v) - \partial_v^{M_2} f(v')| \le A \frac{|v-v'|^\delta}{(1+|v|)^{1+(3/2)M+\gamma+2\delta}} \quad \text{for all } |v-v'| \le \tfrac{1}{2}(1+|v|).$$

For $f \in \mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})$, denote the realization of f as an X-valued map by $\tilde{f} : \mathbb{H}^n \to \mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{R})$. Then, from (5-7) and Theorem 47, we have

$$\|Tf\|_{\mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})} = \|T\tilde{f}\|_{\mathcal{M}^{M'+\delta}(\mathbb{H}^n; \mathcal{M}^{M'+\delta}(\mathbb{R}))} \le C\|\tilde{f}\|_{\mathcal{M}^{M'+\delta}(\mathbb{H}^n; \mathcal{M}^{M'+\delta}(\mathbb{R}))} = C\|f\|_{\mathcal{M}_{\text{product}}^{M'+\delta}(\mathbb{H}^n \times \mathbb{R})}.$$

This completes the proof of Theorem 44. $\square$

5.3. *Orthogonality estimates and the proof of the Plancherel–Pólya inequalities.* We will need *almost-orthogonality estimates* in order to prove both the Plancherel–Pólya inequalities and the boundedness of flag singular integrals on $H_{\text{flag}}^p(\mathbb{H}^n)$. Recall from (2-2) the definition of the components $\psi_{t,s}$ of the continuous decomposition of the identity adapted to the Heisenberg group:

$$\psi(z, u) = \psi^{(1)} *_2 \psi^{(2)}(z, u) = \int_{\mathbb{R}} \psi^{(1)}(z, u-v) \psi^{(2)}(v)\, dv, \quad (z, u) \in \mathbb{C}^n \times \mathbb{R},$$

and

$$\psi_{t,s}(z, u) = \psi_t^{(1)} *_2 \psi_s^{(2)}(z, u) = \int_{\mathbb{R}} \psi_t^{(1)}(z, u-v) \psi_s^{(2)}(v)\, dv = \int_{\mathbb{R}} t^{-2n-2} \psi^{(1)}\left(\frac{z}{t}, \frac{u-v}{t^2}\right) s^{-1} \psi^{(2)}\left(\frac{v}{s}\right) dv.$$

Here $\psi^{(1)} \in \mathscr{S}(\mathbb{H}^n)$ is as in Theorem 2, and $\psi^{(2)} \in \mathscr{S}(\mathbb{R})$ satisfies

$$\int_0^\infty |\widehat{\psi^{(2)}}(t\eta)|^2 \frac{dt}{t} = 1$$

for all $\eta \in \mathbb{R}\backslash\{0\}$, along with the moment conditions

$$\int_{\mathbb{H}^n} z^\alpha u^\beta \psi^{(1)}(z, u)\, dz\, du = 0, \quad |\alpha| + 2\beta \leq M,$$

$$\int_{\mathbb{R}} v^\gamma \psi^{(2)}(v)\, dv = 0, \quad \gamma \geq 0,$$

where M may be fixed arbitrarily large.

In particular, the collection of component functions $\{\psi_{t,s}\}_{t,s>0}$ satisfies

$$
\begin{aligned}
\psi_{t,s} &= \psi_t^{(1)} *_2 \psi_s^{(2)}, \\
\psi_t^{(1)}(z, u) &= t^{-2n-2}\psi^{(1)}\left(\frac{z}{t}, \frac{u-v}{t^2}\right), \\
\psi_s^{(2)}(v) &= s^{-1}\psi^{(2)}\left(\frac{v}{s}\right), \\
\psi^{(1)}(z, u)\psi^{(2)}(v) &\in \mathscr{M}_{\text{product}}^{M+\delta}(\mathbb{H}^n \times \mathbb{R}).
\end{aligned}
\tag{5-8}
$$

Of course the conditions in (5-8) imply that $\psi_{t,s} \in \mathscr{M}_{\text{flag}}^M(\mathbb{H}^n)$ for all $t, s > 0$, but (5-8) also contains the implicit dilation information that cannot be expressed solely in terms of $\psi_{1,1}$. Motivated by these considerations we make the following definition.

Definition 49. To each function $\Psi \in \mathscr{M}_{\text{product}}^{M+\delta}(\mathbb{H}^n \times \mathbb{R})$ we associate a collection of *product dilations* $\{\Psi_{t,s}\}_{t,s>0}$ defined by

$$\Psi_{t,s}(z, u, v) = t^{-2n-2}s^{-1}\Psi\left(\left(\frac{z}{t}, \frac{u}{t^2}\right), \frac{v}{s}\right),$$

and a collection of *component functions* $\{\psi_{t,s}\}_{t,s>0}$ defined by

$$\psi_{t,s}(z, u) = \pi\Psi_{t,s}(z, u) = \int_{\mathbb{R}} t^{-2n-2}s^{-1}\Psi\left(\left(\frac{z}{t}, \frac{u-v}{t^2}\right), \frac{v}{s}\right) dv, \quad t, s > 0.$$

Given two functions in $\mathscr{M}_{\text{product}}^{M+\delta}(\mathbb{H}^n \times \mathbb{R})$ and their corresponding collections of component functions we have the *almost-orthogonality* estimates given below. We use $*_{\mathbb{H}^n}$ to denote convolution on the Heisenberg group $\mathbb{H}^n$, and $*_{\mathbb{H}^n \times \mathbb{R}}$ to denote convolution on the product group $\mathbb{H}^n \times \mathbb{R}$. From Lemma 12 we obtain that π intertwines these two convolutions, which we record here.

Lemma 50. *For $\psi_{t,s}$, $\Psi_{t,s}$, $\phi_{t',s'}$, $\Phi_{t',s'}$ as above, we have*

$$\psi_{t,s} *_{\mathbb{H}^n} \phi_{t',s'} = \pi\{\Psi_{t,s} *_{\mathbb{H}^n \times \mathbb{R}} \Phi_{t',s'}\}.\tag{5-9}$$

We now give the orthogonality estimates, first in the product case and then in the flag case. The product case in Lemma 51 will prove crucial in establishing Theorem 41 for the flag molecular space $\mathscr{M}_{\text{flag}}^{M'+\delta}(\mathbb{H}^n)$.

For convenience, we give the almost orthogonal estimates only for the case $\mathscr{M}_{\text{product}}^{4M+2,2M,2M}(\mathbb{H}^n \times \mathbb{R})$.

Lemma 51. *Suppose* $\Psi, \Phi \in \mathcal{M}_{\text{product}}^{4M+2,2M,2M}(\mathbb{H}^n \times \mathbb{R})$. *Then there exists a constant* $C = C(M)$ *depending only on* M *such that*

$$|\Psi_{t,s} *_{\mathbb{H}^n \times \mathbb{R}} \Phi_{t',s'}((z,u),v)|$$
$$\leq C\left(\frac{t}{t'} \wedge \frac{t'}{t}\right)^{2M+1} \left(\frac{s}{s'} \wedge \frac{s'}{s}\right)^{M+1} \frac{(t \vee t')^{2(4M+2)/2}}{((t \vee t')^2 + |z|^2 + |u|)^{(Q+4M+2)/2}} \frac{(s \vee s')^{4M+2}}{(s \vee s' + |v|)^{1+4M+2}}. \quad (5\text{-}10)$$

The proof of Lemma 51 uses a standard orthogonality argument on the integral

$$\Psi_{t,s} *_{\mathbb{H}^n \times \mathbb{R}} \Phi_{t',s'}((z,u),v) = \int_{\mathbb{H}^n \times \mathbb{R}} \Psi_{t,s}((z,u) \circ (z',u')^{-1}, v - v') \Phi_{t',s'}((z',u'),v') \, dz' \, du' \, dv', \quad (5\text{-}11)$$

and we refer the reader to [Han et al. 2012] for details.

There are corresponding orthogonality estimates for component functions on $\mathbb{H}^n$.

Lemma 52. *Suppose* $\Psi, \Phi \in \mathcal{M}_{\text{product}}^{2M}(\mathbb{H}^n \times \mathbb{R})$ *and let* $\{\psi_{t,s}\}_{t,s>0}$ *and* $\{\phi_{t,s}\}_{t,s>0}$ *be the associated collections of component functions as defined in* (2-7) *above. Then there exists a constant* $C = C(M)$ *depending only on* M *such that, if* $(t \vee t')^2 \leq s \vee s'$, *then*

$$|\psi_{t,s} *_{\mathbb{H}^n} \phi_{t',s'}(z,u)| \leq C\left(\frac{t}{t'} \wedge \frac{t'}{t}\right)^{2M} \left(\frac{s}{s'} \wedge \frac{s'}{s}\right)^{M} \times \frac{(t \vee t')^{2M}}{(t \vee t' + |z|)^{2n+2M}} \frac{(s \vee s')^{M}}{(s \vee s' + |u|)^{1+M}}, \quad (5\text{-}12)$$

and if $(t \vee t')^2 \geq s \vee s'$, *then*

$$|\psi_{t,s} * \phi_{t',s'}(z,u)| \leq C\left(\frac{t}{t'} \wedge \frac{t'}{t}\right)^{M} \left(\frac{s}{s'} \wedge \frac{s'}{s}\right)^{M} \times \frac{(t \vee t')^{M}}{(t \vee t' + |z|)^{2n+M}} \frac{(t \vee t')^{M}}{(t \vee t' + \sqrt{|u|})^{2+2M}}. \quad (5\text{-}13)$$

Roughly speaking, $\psi_{t,s} * \phi_{t',s'}(z,u)$ satisfies the *product multiparameter* almost-orthogonality when $(t \vee t')^2 \leq s \vee s'$ and the *one-parameter* almost-orthogonality when $(t \vee t')^2 \geq s \vee s'$.

Proof of Lemma 52. We will use Lemma 50 to pass from the orthogonality estimates for the product dilations $\{\Psi_{t,s}\}_{t,s>0}$ and $\{\Phi_{t,s}\}_{t,s>0}$ in Lemma 51 to the estimates for the component functions $\{\psi_{t,s}\}_{t,s>0}$ and $\{\phi_{t,s}\}_{t,s>0}$ in Lemma 52.

From (5-10) and (5-9) we obtain

$$|\psi_{t,s} * \phi_{t',s'}(z,u)|$$
$$= \left|\int_{\mathbb{R}} \Psi_{t,s} *_{\mathbb{H}^n \times \mathbb{R}} \Phi_{t',s'}((z,u-v),v) \, dv\right|$$
$$\lesssim C\left(\frac{t}{t'} \wedge \frac{t'}{t}\right)^{2M} \left(\frac{s}{s'} \wedge \frac{s'}{s}\right)^{M} \times \int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u-v|)^{n+1+2M}} \frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} \, dv. \quad (5\text{-}14)$$

Now we consider four cases separately.

Case 1: $(t \vee t')^2 \leq s \vee s'$ *and* $|u| \geq s \vee s'$. In this case we use the fact that

$$\frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} = \frac{1}{s \vee s'} \frac{1}{(1 + |v|/(s \vee s')|)^{1+2M}} \quad (5\text{-}15)$$

has integral roughly 1, with essential support $[-s \vee s', s \vee s']$, to obtain

$$\int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u - v|)^{n+1+2M}} \frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} \, dv$$

$$\approx \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u|)^{n+1+2M}} \leq \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |z|^2)^{n+M}} \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |u|)^{1+M}}$$

$$\leq \frac{(t \vee t')^{2M}}{((t \vee t') + |z|)^{2n+2M}} \frac{(s \vee s')^{M}}{(s \vee s' + |u|)^{1+M}}.$$

Plugging this estimate into the right side of (5-14) leads to the correct product estimate (5-12) for this case.

Case 2: $(t \vee t')^2 \leq s \vee s'$ *and* $|u| \leq s \vee s'$. In this case we bound the left side of (5-15) by $1/(s \vee s')$ to obtain

$$\int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u - v|)^{n+1+2M}} \frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} \, dv$$

$$\lesssim \frac{1}{s \vee s'} \int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u - v|)^{n+1+2M}} \, dv$$

$$\lesssim \frac{1}{s \vee s'} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2)^{n+2M}} \leq \frac{(t \vee t')^{2M}}{((t \vee t') + |z|)^{2n+2M}} \frac{(s \vee s')^{M}}{(s \vee s' + |u|)^{1+M}}.$$

Plugging this estimate into the right side of (5-14) again leads to the correct product estimate (5-12) for this case.

Case 3: $(t \vee t')^2 \geq s \vee s'$ *and* $|u| \leq (t \vee t')^2$. In this case we have

$$\int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u - v|)^{n+1+2M}} \frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} \, dv$$

$$\lesssim \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2)^{n+1+2M}} \lesssim \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |z|^2)^{n+M}} \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |u|)^{1+M}}$$

$$\approx \frac{(t \vee t')^{2M}}{(t \vee t' + |z|)^{2n+2M}} \frac{(t \vee t')^{2M}}{(t \vee t' + \sqrt{|u|})^{2+2M}}.$$

Plugging this estimate into the right side of (5-14) leads to the correct one-parameter estimate (5-13) for this case.

Case 4: $(t \vee t')^2 \geq s \vee s'$ *and* $|u| \geq (t \vee t')^2$. In this case we have

$$\int_{\mathbb{R}} \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u - v|)^{n+1+2M}} \frac{(s \vee s')^{2M}}{(s \vee s' + |v|)^{1+2M}} \, dv$$

$$\lesssim \frac{(t \vee t')^{4M}}{((t \vee t')^2 + |z|^2 + |u|)^{n+1+2M}} \lesssim \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |z|^2)^{n+M}} \frac{(t \vee t')^{2M}}{((t \vee t')^2 + |u|)^{1+M}}$$

$$\approx \frac{(t \vee t')^{2M}}{(t \vee t' + |z|)^{2n+2M}} \frac{(t \vee t')^{2M}}{(t \vee t' + \sqrt{|u|})^{2+2M}}.$$

Plugging this estimate into the right side of (5-14) again leads to the correct one-parameter estimate (5-13). $\qquad\square$

5.3.1. *Proof of the Plancherel–Pólya inequalities.* Before we prove the Plancherel–Pólya-type inequality in Theorem 19, we first prove the following lemma. We will often use the notation (x_I, y_J) in place of (z_I, u_J) for the center of the dyadic rectangle $I \times J$ in $\mathbb{H}^n$; that is, we write x in place of z, and y in place of u.

Lemma 53. *Let $I \times J$ and $I' \times J'$ be dyadic rectangles in $\mathbb{H}^n$ such that*

$$\ell(I) = 2^{-j-N}, \quad \ell(J) = 2^{-j-N} + 2^{-k-N}, \quad \ell(I') = 2^{-j'-N}, \quad \text{and } \ell(J') = 2^{-j'-N} + 2^{-k'-N}.$$

Thus, for any (u, v) and (u^*, v^*) in $\mathbb{H}^n$, we have, when $j \wedge j' \geq k \wedge k'$,

$$\sum_{I', J'} \frac{2^{-|j-j'|L_1 - |k-k'|L_2} 2^{-(j \wedge j')K_1 - (k \wedge k')K_2} |I'||J'|}{(2^{-j \wedge j'} + |u - x_{I'}|)^{2n+K_1} (2^{-k \wedge k'} + |v - y_{J'}|)^{1+K_2}} \cdot |\phi_{j',k'} * f(x_{I'}, y_{J'})|$$

$$\leq C_1(N, r, j, j', k, k') 2^{-|j-j'|L_1} \times 2^{-|k-k'|L_2} \left\{ M_S\left[\left(\sum_{J'} \sum_{I'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right] \right\}^{\frac{1}{r}} (u^*, v^*),$$

and when $j \wedge j' \leq k \wedge k'$,

$$\sum_{I', J'} \frac{2^{-|j-j'|L_1 - |k-k'|L_2} 2^{-(j \wedge j')K_1 - (j \wedge j')K_2} |I'||J'|}{(2^{-j \wedge j'} + |u - x_{I'}|)^{2n+K_1} (2^{-j \wedge j'} + |v - y_{J'}|)^{1+K_2}} |\phi_{j',k'} * f(x_{I'}, y_{J'})|$$

$$\leq C_2(N, r, j, j', k, k') 2^{-|j-j'|L_1} 2^{-|k-k'|L_2} \times \left\{ M\left[\left(\sum_{J'} \sum_{I'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right] \right\}^{\frac{1}{r}} (u^*, v^*),$$

where M is the Hardy–Littlewood maximal function on $\mathbb{H}^n$, M_S is the strong maximal function on $\mathbb{H}^n$ as defined in (1-1), $\max\{2n/(2n + K_1), 1/(1 + K_2)\} < r$ and

$$C_1(N, r, j, j', k, k') = 2^{(1/r-1)N(2n+1)} \cdot 2^{[2n(j \wedge j' - j') + (k \wedge k' - k')](1-1/r)},$$
$$C_2(N, r, j, j', k, k') = 2^{(1/r-1)N(2n+1)} \cdot 2^{[2n(j \wedge j' - j') + (j \wedge j' - j' \wedge k')](1-1/r)}.$$

Proof. We set

$$A_0 = \left\{ I' : \ell(I') = 2^{-j'-N}, \ \frac{|u - x_{I'}|}{2^{-j \wedge j'}} \leq 1 \right\},$$

$$B_0 = \left\{ J' : \ell(J') = 2^{-j'-N} + 2^{-k'-N}, \ \frac{|v - y_{J'}|}{2^{-k \wedge k'}} \leq 1 \right\},$$

where $x_{I'} \in I'$ and $y_{J'} \in J'$, and where, for $\ell \geq 1$, $i \geq 1$,

$$A_\ell = \left\{ I' : \ell(I') = 2^{-j'-N}, \ 2^{\ell-1} < \frac{|u - x_{I'}|}{2^{-j \wedge j'}} \leq 2^\ell \right\}.$$

$$B_i = \left\{ J' : \ell(J') = 2^{-j'-N} + 2^{-k'-N}, \ 2^{i-1} < \frac{|v - y_{J'}|}{2^{-k \wedge k'}} \leq 2^i \right\}.$$

We first consider the case when $j \wedge j' \geq k \wedge k'$, and let

$$\tau = [2n(j \wedge j' - j') + (k \wedge k' - k')]\left(1 - \frac{1}{r}\right).$$

Then

$$\sum_{I',J'} \frac{2^{-(j\wedge j')K_1 - (k\wedge k')K_2}|I'||J'|}{(2^{-j\wedge j'} + |u - x_{I'}|)^{2n+K_1}(2^{-k\wedge k'} + |v - y_{J'}|)^{1+K_2}} \cdot |\phi_{j',k'} * f(x_{I'}, y_{J'})|$$

$$\leq \sum_{\ell, i \geq 0} 2^{-\ell(2n+K_1)}2^{-i(1+K_2)}2^{-N(2n+1)}2^{(j\wedge j'-j')n+(k\wedge k'-k')m} \sum_{I' \in A_\ell, J' \in B_i} |\phi_{j',k'} * f(x_{I'}, y_{J'})|$$

$$\leq \sum_{\ell, i \geq 0} 2^{-\ell(n+K_1)}2^{-i(m+K_2)}2^{-N(n+m)}2^{(j\wedge j'-j')2n+(k\wedge k'-k')}\left(\sum_{I' \in A_\ell, J' \in B_i} (|\phi_{j',k'} * f(x_{I'}, y_{J'})|)^r\right)^{1/r}$$

$$= \sum_{\ell, i \geq 0} 2^{-\ell(2n+K_1)-i(1+K_2)-N(2n+1)}2^{(j\wedge j'-j')2n+(k\wedge k'-k')}$$

$$\times \left(\int_{\mathbb{H}^n} |I'|^{-1}|J'|^{-1} \sum_{I' \in A_\ell, J' \in B_i} |\phi_{j',k'} * f(x_{I'}, y_{J'})|^r \chi_{I'}\chi_{J'}\right)^{1/r}$$

$$\leq \sum_{\ell, i \geq 0} 2^{-\ell(2n+K_1-2n/r)-i(1+K_2-1/r)+(1/r-1)N(2n+1)}$$

$$\times 2^\tau \left(M_S\left(\sum_{I' \in A_\ell, J' \in B_i} |\phi_{j',k'} * f(x_{I'}, y_{J'})|^r \chi_{I'}\chi_{J'}\right)(u^*, v^*)\right)^{1/r}$$

$$\leq C_1(N, r, j, k, j', k')\left(M_S\left(\sum_{I',J'} |\phi_{j',k'} * f(x_{I'}, y_{J'})|^r \chi_{I'}\chi_{J'}\right)(u^*, v^*)\right)^{1/r}$$

The last inequality follows from the assumption that $r > \max\{2n/(2n + K_1), 1/(1 + K_2)\}$, which can be achieved by choosing K_1, K_2 large enough. The second inequality can be proved similarly. $\square$

We are now ready to give the proof of the Plancherel–Pólya inequality.

Proof of Theorem 19. By Theorem 17, $f \in \mathcal{M}_{\text{flag}}^{M+\delta}(\mathbb{H}^n)'$ can be represented by

$$f(z, u) = \sum_{j'} \sum_{k'} \sum_{J'} \sum_{I'} |J'||I'|\tilde{\phi}_{j',k'}((z, u) \circ (x_{I'}, y_{J'})^{-1})(\phi_{j',k'} * f)(x_{I'}, y_{J'}).$$

We write

$$(\psi_{j,k} * f)(u, v) = \sum_{j'} \sum_{k'} \sum_{J'} \sum_{I'} |I'||J'|(\psi_{j,k} * \tilde{\phi}_{j',k'}((\cdot, \cdot) \circ (x_{I'}, y_{J'})^{-1}))(u, v)(\phi_{j',k'} * f)(x_{I'}, y_{J'}).$$

By the almost-orthogonality estimates in Lemma 52, and by choosing $t = 2^{-j}$, $s = 2^{-k}$, $t' = 2^{-j'}$, $s' = 2^{-k'}$, and for any given positive integers L_1, L_2, K_1, K_2, we have, if $j \wedge j' \geq k \wedge k'$,

$$|(\psi_{j,k} * \tilde{\phi}_{j',k'}((\cdot, \cdot) \circ (x_{I'}, y_{J'})^{-1}))(u, v)|$$

$$\leq \frac{2^{-|j-j'|L_1-|k-k'|L_2}2^{-(j\wedge j')K_1-(k\wedge k')K_2}|I'||J'|}{(2^{-j\wedge j'} + |u - x_{I'}|)^{2n+K_1}(2^{-k\wedge k'} + |v - y_{J'}|)^{1+K_2}}|\phi_{j',k'} * f(x_{I'}, y_{J'})|,$$

and when $j \wedge j' \leq k \wedge k'$, we have

$$|(\psi_{j,k} * \tilde{\phi}_{j',k'}((\cdot, \cdot) \circ (x_{I'}, y_{J'})^{-1}))(u, v)|$$

$$\leq \frac{2^{-|j-j'|L_1 - |k-k'|L_2} 2^{-(j \wedge j')K_1 - (j \wedge j')K_2} |I'||J'|}{(2^{-j \wedge j'} + |u - x_{I'}|)^{2n+K_1}(2^{-j \wedge j'} + |v - y_{J'}|)^{1+K_2}} |\phi_{j',k'} * f(x_{I'}, y_{J'})|.$$

Using Lemma 53, for any $u, u^* \in I$, $x_{I'} \in I'$, $v, v^* \in J$, and $y_{J'} \in J'$, we have

$$|\psi_{j,k} * f(u, v)|$$

$$\leq C_1 \sum_{j',k':j \wedge j' \geq k \wedge k'} 2^{-|j-j'|L_1} \cdot 2^{-|k-k'|L_2} \times \left\{ M_S \left[\left(\sum_{J'} \sum_{I'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right] \right\}^{1/r} (u^*, v^*)$$

$$+ C_2 \sum_{j',k':j \wedge j' \leq k \wedge k'} 2^{-|j-j'|L_1} \cdot 2^{-|k-k'|L_2} \times \left\{ M \left[\left(\sum_{J'} \sum_{I'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right] \right\}^{1/r} (u^*, v^*)$$

$$\leq C \sum_{j',k'} 2^{-|j-j'|L_1} \cdot 2^{-|k-k'|L_2} \times \left\{ M_S \left[\left(\sum_{J'} \sum_{I'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right] \right\}^{1/r} (u^*, v^*),$$

where M is the Hardy–Littlewood maximal function on $\mathbb{H}^n$, M_S is the strong maximal function on $\mathbb{H}^n$, and $\max\{2n/(2n + K_1), 1/(1 + K_2)\} < r < p$.

Applying Hölder's inequality and summing over j, k, I, J yields

$$\left\{ \sum_{j,k} \sum_{I,J} \sup_{u \in I, v \in J} |\psi_{j,k} * f(u, v)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \leq C \left\{ \sum_{j',k'} \left\{ M_S \left(\sum_{I',J'} |\phi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{I'} \chi_{J'} \right)^r \right\}^{2/r} \right\}^{\frac{1}{2}}.$$

Since $x_{I'}$ and $y_{J'}$ are arbitrary points in I' and J', respectively, we have

$$\left\{ \sum_{j,k} \sum_{I,J} \sup_{u \in I, v \in J} |\psi_{j,k} * f(u, v)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \leq C \left\{ \sum_{j',k'} \left\{ M_S \left(\sum_{I',J'} \inf_{\substack{u \in I' \\ v \in J'}} |\phi_{j',k'} * f(u, v)| \chi_{I'} \chi_{J'} \right)^r \right\}^{2/r} \right\}^{\frac{1}{2}},$$

and hence, by the Fefferman–Stein vector-valued maximal function inequality [Fefferman and Stein 1982] with $r < p$, we get

$$\left\| \left\{ \sum_j \sum_k \sum_J \sum_I \sup_{\substack{u \in I \\ v \in J}} |\psi_{j,k} * f(u, v)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_p \leq C \left\| \left\{ \sum_{j'} \sum_{k'} \sum_{J'} \sum_{I'} \inf_{\substack{u \in I' \\ v \in J'}} |\phi_{j',k'} * f(u, v)|^2 \chi_{I'} \chi_{J'} \right\}^{\frac{1}{2}} \right\|_p.$$

This completes the proof of Theorem 19. □

6. Boundedness of flag singular integrals

As a consequence of Theorem 19, it is easy to see that the Hardy space H^p_{flag} is independent of the choice of the functions ψ. Moreover, we have the following characterization of H^p_{flag} using the wavelet norm.

Proposition 54. *Let $0 < p \leq 1$. Then we have*

$$\|f\|_{H^p_{\text{flag}}} \approx \left\| \left\{ \sum_j \sum_k \sum_J \sum_I |\psi_{j,k} * f(x_I, y_J)|^2 \chi_I(x) \chi_J(y) \right\}^{\frac{1}{2}} \right\|_p,$$

where $j, k, \psi, \chi_I, \chi_J, x_I, y_J$ are as in Theorem 19.

Before we give the proof of the boundedness of flag singular integrals on H^p_{flag}, we demonstrate several properties of H^p_{flag}.

Proposition 55. $\mathcal{M}^{M+\delta}_{\text{flag}}(\mathbb{H}^n)$ *is dense in* $H^p_{\text{flag}}(\mathbb{H}^n)$ *for M large enough.*

Proof. Suppose $f \in H^p_{\text{flag}}$, and set $W = \{(j, k, I, J) : |j| \leq L, |k| \leq M, I \times J \subseteq B(0, r)\}$, where $I \times J$ is a dyadic rectangle in $\mathbb{H}^n$ with $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-k-N} + 2^{-j-N}$, and where $B(0, r)$ is the ball in $\mathbb{H}^n$ centered at the origin with radius r. It is easy to see that

$$\sum_{(j,k,I,J)\in W} |I||J|\tilde{\psi}_{j,k}((z, y) \circ (x_I, y_J)^{-1}) \psi_{j,k} * f(x_I, y_J)$$

is a test function in $\mathcal{M}^{M+\delta}_{\text{flag}}(\mathbb{H}^n)$ for any fixed L, M, r. To obtain the proposition, it suffices to prove

$$\sum_{(j,k,I,J)\in W^c} |I||J|\tilde{\psi}_{j,k}((z, y) \circ (x_I, y_J)^{-1}) \psi_{j,k} * f(x_I, y_J)$$

tends to zero in the H^p_{flag} norm as L, M, r tend to infinity. This follows from an argument similar to that in the proof of Theorem 19. In fact, repeating the argument in Theorem 19 yields

$$\left\| \sum_{(j,k,I,J)\in W^c} |I||J|\tilde{\psi}_{j,k}((z, y) \circ (x_I, y_J)^{-1}) \psi_{j,k} * f(x_I, y_J) \right\|_{H^p_{\text{flag}}}$$

$$\leq C \left\| \left\{ \sum_{(j,k,I,J)\in W^c} |\psi_{j,k} * f(x_I, y_J)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_p,$$

where the last term tends to zero as L, M, r tend to infinity whenever $f \in H^p_{\text{flag}}$. $\qquad\square$

As a consequence of Proposition 55, $L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$ is dense in $H^p_{\text{flag}}(\mathbb{H}^n)$. Furthermore, we have the following theorem.

Theorem 56. *If $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n), 0 < p \leq 1$, then $f \in L^p(\mathbb{H}^n)$ and there is a constant $C_p > 0$ which is independent of the L^2 norm of f such that*

$$\|f\|_p \leq C\|f\|_{H^p_{\text{flag}}}.$$

To prove Theorem 56, we need a discrete Calderón reproducing formula on $L^2(\mathbb{H}^n)$. To be more precise, take $\phi^{(1)} \in C_0^\infty(\mathbb{H}^n)$ as in Theorem 2 with

$$\int_{\mathbb{H}^n} \phi^{(1)}(z, u) z^\alpha u^\beta dz \, du = 0 \quad \text{for all } \alpha, \beta \text{ satisfying } 0 \leq |\alpha| \leq M_0, \ 0 \leq |\beta| \leq M_0,$$

and take $\phi^{(2)} \in C_0^\infty(\mathbb{R})$ with

$$\int_{\mathbb{R}} \phi^{(2)}(v) z^\gamma \, dv = 0 \quad \text{for all } 0 \le |\gamma| \le M_0,$$

and $\sum_k |\widehat{\phi^{(2)}}(2^{-k}\xi_2)|^2 = 1$ for all $\xi_2 \in \mathbb{R} \backslash \{0\}$.

Furthermore, we may assume that $\phi^{(1)}$ and $\phi^{(2)}$ are radial functions and supported in the unit balls of $\mathbb{H}^n$ and $\mathbb{R}$, respectively. Set

$$\phi_{jk}(z, u) = \int_{\mathbb{R}} \phi_j^{(1)}(z, u - v) \phi_k^{(2)}(v) \, dv.$$

By Theorem 2 we have the following continuous version of the Calderón reproducing formula on L^2: for $f \in L^2(\mathbb{H}^n)$,

$$f(z, u) = \sum_j \sum_k \phi_{jk} * \phi_{jk} * f(z, u).$$

For our purposes, we need a discrete version of the above reproducing formula.

Theorem 57. *There exist functions $\tilde{\phi}_{jk}$ and an operator T_N^{-1} such that*

$$f(x, y) = \sum_j \sum_k \sum_J \sum_I |I| |J| \tilde{\phi}_{j,k}((x, y) \circ (x_I, y_J)^{-1}) \phi_{j,k} * (T_N^{-1}(f))(x_I, y_J),$$

where the functions $\tilde{\phi}_{jk}((x, y) \circ (x_I, y_J)^{-1})$ satisfy the conditions in Theorem 17 with $\alpha_1, \beta_1, \gamma_1, N, M$ depending on M_0. Moreover, T_N^{-1} is bounded on both $L^2(\mathbb{H}^n)$ and $H_{\mathrm{flag}}^p(\mathbb{H}^n)$, and the series converges in $L^2(\mathbb{H}^n)$.

Remark 58. The difference between Theorems 57 and 17 is that the $\tilde{\phi}_{jk}$ in Theorem 57 have compact support. The price we pay here is that $\tilde{\phi}_{jk}$ only satisfies moment conditions of finite order, unlike in Theorem 17, where moment conditions of infinite order are satisfied. Moreover, the formula in Theorem 57 only holds on $L^2(\mathbb{H}^n)$ while the formula in Theorem 17 holds in both the test function space $\mathcal{M}_{\mathrm{flag}}^{M+\delta}$ and its dual space $(\mathcal{M}_{\mathrm{flag}}^{M+\delta})'$.

Proof of Theorem 57. Following the proof of Theorem 17, we have

$$f(z, u) = \sum_j \sum_k \sum_J \sum_I \left[\int_J \int_I \phi_{j,k}((z, u) \circ (u, v)^{-1}) \, du \, dv \right] (\phi_{j,k} * f)(x_I, y_J) + \mathcal{R} f(x, y),$$

where I, J, j, k, and $\mathcal{R}$ are as in Theorem 17.

We need the following lemma to handle the remainder term $\mathcal{R}$.

Lemma 59. *Let $0 < p \le 1$. Then the operator $\mathcal{R}$ is bounded on $L^2(\mathbb{H}^n)$ and $H_{\mathrm{flag}}^p(\mathbb{H}^n)$ whenever M_0 is chosen to be a large positive integer. Moreover, there exists a constant $C > 0$ such that*

$$\|\mathcal{R} f\|_2 \le C 2^{-N} \|f\|_2 \quad \text{and} \quad \|\mathcal{R} f\|_{H_{\mathrm{flag}}^p(\mathbb{H}^n)} \le C 2^{-N} \|f\|_{H_{\mathrm{flag}}^p(\mathbb{H}^n)}.$$

Proof. Following the proofs of Theorems 17 and 19 and using the wavelet Calderón reproducing formula for $f \in L^2(\mathbb{H}^n)$, we have

$$\|g_{\text{flag}}(\mathcal{R}f)\|_p$$

$$\leq \left\| \left\{ \sum_j \sum_k \sum_J \sum_I |(\psi_{j,k} * \mathcal{R}f)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_p$$

$$= \left\| \left\{ \sum_{j,k,J,I} \sum_{j',k',J',I'} |J'||I'| |(\psi_{j,k} * \mathcal{R}\tilde{\psi}_{j',k'}((\cdot,\cdot) \circ (x_{I'}, y_{J'})^{-1}) \cdot \psi_{j'k'} * f(x_{I'}, y_{J'}))|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_p,$$

where $j, k, \psi, \chi_I, \chi_J, x_I, y_J$ are as in Theorem 19.

Claim. *We have*

$$|(\psi_{j,k} * \mathcal{R}(\tilde{\psi}_{j',k'}((\cdot,\cdot) \circ (x_{I'}, y_{J'})^{-1})))(z,u)|$$

$$\leq C 2^{-N} 2^{-|j-j'|K} 2^{-|k-k'|K} \int_{\mathbb{R}} \frac{2^{-(j \wedge j')K}}{(2^{-(j \wedge j')} + |z - x_{I'}| + |u - v - y_{J'}|)^{2n+1+K}} \frac{2^{-(k \wedge k')K}}{(2^{-(k \wedge k')} + |v|)^{1+K}} \, dv,$$

where, for simplicity, we have chosen

$$L_1 = L_2 = K_1 = K_2 = K < M_0, \quad \max\left(\frac{2n}{2n+K}, \frac{1}{1+K}\right) < p,$$

and M_0 is chosen to be a larger integer later.

Assuming the claim for the moment, we can repeat an argument used in Lemma 53, and then use Theorem 19 to obtain

$$\|g_{\text{flag}}(\mathcal{R}f)\|_p \leq C 2^{-N} \left\| \left\{ \sum_{j'} \sum_{k'} \left[M_S \left(\sum_{J'} \sum_{I'} |\psi_{j',k'} * f(x_{I'}, y_{J'})| \chi_{J'} \chi_{I'} \right)^r \right]^{2/r} \right\}^{\frac{1}{2}} \right\|_p$$

$$\leq C 2^{-N} \left\| \left\{ \sum_{j'} \sum_{k'} \sum_{J'} \sum_{I'} |\psi_{j',k'} * f(x_{I'}, y_{J'})|^2 \chi_{I'} \chi_{J'} \right\}^{\frac{1}{2}} \right\|_p \leq C 2^{-N} \|f\|_{H^p_{\text{flag}}(\mathbb{H}^n)}.$$

It is clear that the above estimates continue to hold when p is replaced by 2. This completes the proof of Lemma 59 modulo the claim.

In order to prove the claim made above, we note that Theorem 41 shows that the functions

$$\mathcal{R}(\tilde{\psi}_{j',k'}((\cdot,\cdot) \circ (x_{I'}, y_{J'})^{-1}))(z,u)$$

are flag molecules. Then the claim follows from Lemma 53, and this completes the proof of Lemma 59. $\square$

We now return to the proof of Theorem 57. Let $(T_N)^{-1} = \sum_{i=1}^{\infty} \mathcal{R}^i$, where

$$T_N f = \sum_j \sum_k \sum_J \sum_I \left(\frac{1}{|I||J|} \int_J \int_I \phi_{j,k}((x,y) \circ (u,v)^{-1}) \, du \, dv \right) |I||J| (\phi_{j,k} * f)(x_I, y_J).$$

Lemma 59 shows that if N is large enough, then both T_N and $(T_N)^{-1}$ are bounded on $L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$. Hence, we can get the reproducing formula

$$
f(x, y) = \sum_j \sum_k \sum_J \sum_I |I||J| \tilde{\phi}_{j,k}((x, y) \circ (x_I, y_J)^{-1}) \phi_{j,k} * (T_N^{-1} f)(x_I, y_J),
$$

where the functions $\tilde{\phi}_{jk}((x, y) \circ (x_I, y_J)^{-1})$ are flag molecules, and the series converges in $L^2(\mathbb{H}^n)$. This completes the proof of Theorem 57. $\qquad\square$

As a consequence of Theorem 57, we obtain the following corollary.

Corollary 60. *If $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$ and $0 < p \le 1$, then*

$$
\|f\|_{H^p_{\text{flag}}} \approx \left\| \left\{ \sum_j \sum_k \sum_J \sum_I |\phi_{jk} * (T_N^{-1} f)(x_I, y_J)|^2 \chi_I(z) \chi_J(u) \right\}^{\frac{1}{2}} \right\|_p,
$$

where the constants are independent of the L^2 norm of f.

Proof. Note that if $f \in L^2(\mathbb{H}^n)$, we can apply the Calderón reproducing formula in Theorem 57 and then repeat the proof of Theorem 19. We leave the details to the reader. $\qquad\square$

We now start the proof of Theorem 56. We define a square function by

$$
\tilde{g}(f)(z, u) = \left\{ \sum_j \sum_k \sum_J \sum_I |\phi_{j,k} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I(z) \chi_J(u) \right\}^{\frac{1}{2}},
$$

where the ϕ_{jk} are as in Theorem 57. By Corollary 60, for $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$, we have

$$
\|\tilde{g}(f)\|_{L^p(\mathbb{H}^n)} \le C \|f\|_{H^p_{\text{flag}}(\mathbb{H}^n)}.
$$

To complete the proof of Theorem 56, let $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$. Set

$$
\Omega_i = \{(z, u) \in \mathbb{H}^n : \tilde{g}(f)(z, u) > 2^i\}.
$$

Let

$$
\mathcal{B}_i = \{(j, k, I, J) : |(I \times J) \cap \Omega_i| > \tfrac{1}{2}|I \times J|, |(I \times J) \cap \Omega_{i+1}| \le \tfrac{1}{2}|I \times J|\},
$$

where $I \times J$ are rectangles in $\mathbb{H}^n$ with side lengths $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-k-N} + 2^{-j-N}$. Since $f \in L^2(\mathbb{H}^n)$, the discrete Calderón reproducing formula in Theorem 57 gives

$$
f(z, u) = \sum_j \sum_k \sum_J \sum_I \tilde{\phi}_{j,k}((z, u) \circ (x_I, y_J)^{-1}) |I||J| \phi_{j,k} * (T_N^{-1}(f))(x_I, y_J)
$$

$$
= \sum_i \sum_{(j,k,I,J) \in \mathcal{B}_i} |I||J| \tilde{\phi}_{j,k}((z, u) \circ (x_I, y_J)^{-1}) \phi_{j,k} * (T_N^{-1}(f))(x_I, y_J),
$$

where the series converges rapidly in L^2 norm, and hence almost everywhere.

Claim. *We have*

$$\left\| \sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J|\tilde{\phi}_{j,k}((z,u)\circ(x_I,y_J)^{-1})\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J) \right\|_p^p \leq C2^{ip}|\Omega_i|,$$

which together with the fact that $0 < p \leq 1$ yields

$$\|f\|_p^p \leq \sum_i \left\| \sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J|\tilde{\phi}_{j,k}((z,u)\circ(x_I,y_J)^{-1})\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J) \right\|_p^p$$

$$\leq C \sum_i 2^{ip}|\Omega_i| \leq C\|\tilde{g}(f)\|_p^p \leq C\|f\|_{H_{\mathrm{flag}}^p}^p.$$

To obtain the claim, note that $\phi^{(1)}$ and $\psi^{(2)}$ are radial functions supported in unit balls in $\mathbb{H}^n$ and $\mathbb{R}$, respectively. Hence, if $(j,k,I,J) \in \mathcal{B}_i$, then $\phi_{j,k}((z,u)\circ(x_I,y_J)^{-1})$ is supported in

$$\widetilde{\Omega}_i = \left\{ (z,u) : M_S(\chi_{\Omega_i})(z,u) > \tfrac{1}{100} \right\}.$$

Thus, by Hölder's inequality,

$$\left\| \sum_{(j,k,I,J)\in\mathcal{B}_i} |J||I|\tilde{\phi}_{j,k}((z,u)\circ(x_I,y_J)^{-1})\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J) \right\|_p^p$$

$$\leq |\widetilde{\Omega}_i|^{1-p/2} \left\| \sum_{(j,k,I,J)\in\mathcal{B}_i} |J||I|\tilde{\phi}_{j,k}((z,u)\circ(x_I,y_J)^{-1})\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J) \right\|_2^p.$$

By duality, for all $g \in L^2$ with $\|g\|_2 \leq 1$,

$$\left| \left\langle \sum_{(j,k,I,J)\in\mathcal{B}_i} |J||I|\tilde{\phi}_{j,k}((z,u)\circ(x_I,y_J)^{-1})\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J), g \right\rangle \right|$$

$$= \left| \sum_{(j,k,I,J)\in\mathcal{B}_i} |J||I|\tilde{\phi}_{j,k}*g(x_I,y_J)\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J) \right|$$

$$\leq C \left(\sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J||\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J)|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J||\tilde{\phi}_{j,k}*g(x_I,y_J)|^2 \right)^{\frac{1}{2}}.$$

Since

$$\left(\sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J||\tilde{\phi}_{j,k}*g(x_I,y_J)|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{(j,k,I,J)\in\mathcal{B}_i} |I||J|(M_S(\tilde{\phi}_{j,k}*g)(z,u)\chi_I(z)\chi_J(u))^2 \right)^{\frac{1}{2}}$$

$$\leq C \left(\sum_{j,k} \int_{\mathbb{C}^n} \int_{\mathbb{R}} (M_S(\tilde{\phi}_{j,k}*g)^2(z,u)\,dz\,du \right)^{\frac{1}{2}} \leq C\|g\|_2,$$

the claim now follows from the fact that $|\tilde{\Omega}_i| \leq C|\Omega_i|$ and the estimate

$$C2^{2i}|\Omega_i| \geq \int_{\tilde{\Omega}_i \backslash \Omega_{i+1}} \tilde{g}^2(f)(z,u)\,dz\,du \geq \sum_{(j,k,I,J)\in\mathscr{B}_i} |\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J)|^2 |(I\times J)\cap \tilde{\Omega}_i\backslash\Omega_{i+1}|$$

$$\geq \frac{1}{2}\sum_{(j,k,I,J)\in\mathscr{B}_i} |I||J||\phi_{j,k}*(T_N^{-1}(f))(x_I,y_J)|^2,$$

where the fact that $|(I\times J)\cap\tilde{\Omega}_i\backslash\Omega_{i+1}| > \frac{1}{2}|I\times J|$ when $(j,k,I,J)\in\mathscr{B}_i$ is used in the last inequality. This finishes the proof of Theorem 56.

As a consequence of Theorem 56, we have the following corollary.

Corollary 61. $H_{\text{flag}}^1(\mathbb{H}^n)$ *is a subspace of* $L^1(\mathbb{H}^n)$.

Proof. Given $f \in H_{\text{flag}}^1(\mathbb{H}^n)$, by Proposition 55, there is a sequence $\{f_n\}$ such that $f_n \in L^2(\mathbb{H}^n)\cap H_{\text{flag}}^1(\mathbb{H}^n)$ and f_n converges to f in the norm of $H_{\text{flag}}^1(\mathbb{H}^n)$. By Theorem 56, f_n converges to g in $L^1(\mathbb{H}^n)$ for some $g \in L^1(\mathbb{H}^n)$. Therefore, $f = g$ in $(\mathcal{M}_{\text{flag}}^{M+\delta})'$. $\qquad\square$

Proof of Theorem 22. We assume that K is the kernel of T. Applying the discrete Calderón reproducing formula in Theorem 57 implies that, for $f \in L^2(\mathbb{H}^n)\cap H_{\text{flag}}^p(\mathbb{H}^n)$,

$$\left\|\left\{\sum_{j,k}\sum_{I,J}|\phi_{j,k}*K*f(z,u)|^2\chi_I(x)\chi_J(y)\right\}^{\frac{1}{2}}\right\|_p$$

$$= \left\|\left\{\sum_{j,k}\sum_{I,J}\left|\sum_{j',k'}\sum_{I',J'}|J'||I'|\phi_{j,k}*K*\tilde{\phi}_{j',k'}((\,\cdot\,,\cdot\,)\circ(x_I,y_J)^{-1})(z,u)\right.\right.\right.$$
$$\left.\left.\left.\times \phi_{j',k'}*(T_N^{-1}(f))(x_{I'},y_{J'})\right|^2\chi_I(x)\chi_J(y)\right\}^{\frac{1}{2}}\right\|_p,$$

where the discrete Calderón reproducing formula in $L^2(\mathbb{H}^n)$ is used.

Note that the ϕ_{jk} are dilations of bump functions, and by estimates similar to those in (5-10), one can easily check that

$$|\phi_{j,k}*K*\tilde{\phi}_{j',k'}((\,\cdot\,,\cdot\,)\circ(x_{I'},y_{J'})^{-1})(z,u)|$$

$$\leq C2^{-|j-j'|K}2^{-|k-k'|K}\int_{\mathbb{R}} \frac{2^{-(j\wedge j')K}}{(2^{-(j\wedge j')}+|z-x_{I'}|+|u-v-y_{J'}|)^{2n+1+K}} \cdot \frac{2^{-(k\wedge k')K}}{(2^{-(k\wedge k')}+|v|)^{1+K}}\,dv,$$

where K depends on M_0 given in Theorem 22, and M_0 is chosen large enough.

Repeating an argument similar to that in the proof of Theorem 19, together with Corollary 60, we obtain

$$\|Tf\|_{H_{\text{flag}}^p} \leq C\left\|\left\{\sum_{j'}\sum_{k'}\left\{M_S\left(\sum_{J'}\sum_{I'}|\phi_{j',k'}*(T_N^{-1}(f))(x_{I'},y_{J'})|\chi_{J'}\chi_{I'}\right)^r\right\}^{2/r}(z,u)\right\}^{\frac{1}{2}}\right\|_p$$

$$\leq C\left\|\left\{\sum_{j'}\sum_{k'}\sum_{J'}\sum_{I'}|\phi_{j',k'}*(T_N^{-1}(f))(x_{I'},y_{J'})|^2\chi_{J'}(y)\chi_{I'}(x)\right\}^{\frac{1}{2}}\right\|_p \leq C\|f\|_{H_{\text{flag}}^p},$$

where the last inequality follows from Corollary 60.

Since $L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$ is dense in $H^p_{\text{flag}}(\mathbb{H}^n)$, T can be extended to a bounded operator on $H^p_{\text{flag}}(\mathbb{H}^n)$, and this ends the proof of Theorem 22. $\qquad\square$

Proof of Theorem 23. We note that $H^p_{\text{flag}} \cap L^2$ is dense in H^p_{flag}, so we only have to obtain the required inequality for $f \in H^p_{\text{flag}} \cap L^2$. Thus Theorem 23 follows immediately from Theorems 22 and 56. $\qquad\square$

7. Duality of Hardy spaces H^p_{flag}

Chang and Fefferman [1985] established that the dual space of $H^1(\mathbb{H}^n)$ is BMO($\mathbb{H}^n$) by using the bi-Hilbert transform, and, consequently, their method is not directly applicable to the implicit two-parameter structure associated to flag singular integrals. In order to deal with the duality theory of $H^p_{\text{flag}}(\mathbb{H}^n)$ for all $0 < p \leq 1$, we proceed differently, and first prove Theorem 30, the Plancherel–Pólya inequalities for the Carleson space $\text{CMO}^p_{\text{flag}}$. This theorem implies that the function space $\text{CMO}^p_{\text{flag}}$ is well defined.

Proof of Theorem 30. The idea of the proof of this theorem is, as in the proof of Theorem 19, to use the wavelet Calderón reproducing formula and the almost-orthogonality estimate. For convenience, we prove Theorem 30 for the smallest Heisenberg group $\mathbb{H}^1 = \mathbb{C} \times \mathbb{R}$. However, it will be clear from the proof that its extension to general $\mathbb{H}^n$ is straightforward. Moreover, to simplify notation, we denote $f_{j,k} = f_R$, where $R = I \times J \subset \mathbb{H}^1$, $\ell(I) = 2^{-j-N}$, $\ell(J) = 2^{-k-N} + 2^{-j-N}$, I is a dyadic cube in $\mathbb{R}^2$ and J is an interval in $\mathbb{R}$. Here N is the same as in Theorem 17. We also denote by $\text{dist}(I, I')$ the distance between intervals I and I',

$$S_R = \sup_{\substack{u \in I \\ v \in J}} |\psi_R * f(u,v)|^2, \quad T_R = \inf_{\substack{u \in I \\ v \in J}} |\phi_R * f(u,v)|^2.$$

With this notation, we can rewrite the wavelet Calderón reproducing formula in Theorem 17 as

$$f(z, u) = \sum_{R=I \times J} |I||J| \tilde{\phi}_R(z, u) \phi_R * f(x_I, y_J),$$

where the sum runs over all rectangles $R = I \times J$. Let

$$R' = I' \times J', \quad |I'| = 2^{-j'-N}, \quad |J'| = 2^{-j'-N} + 2^{-k'-N}.$$

Applying the above wavelet Calderón reproducing formula and the almost-orthogonality estimates in Section 5.3 yields, for all $(u, v) \in R$,

$$
\begin{aligned}
|\psi_R * f(u,v)|^2 \leq C &\sum_{\substack{R'=I' \times J' \\ j' > k'}} \left(\frac{|I|}{|I'|} \wedge \frac{|I'|}{|I|} \right)^L \left(\frac{|J|}{|J'|} \wedge \frac{|J'|}{|J|} \right)^L \\
&\times \frac{|I'|^K}{(|I'| + |u - x_{I'}|)^{(1+K)}} \frac{|J'|^K}{(|J'| + |v - y_{J'}|)^{(1+K)}} |I'||J'| |\phi_{R'} * f(x_{I'}, y_{J'})|^2 \\
+ C &\sum_{\substack{R'=I'J' \\ j' \leq k'}} \left(\frac{|I|}{|I'|} \wedge \frac{|I'|}{|I|} \right)^L \left(\frac{|J|}{|J'|} \wedge \frac{|J'|}{|J|} \right)^L \\
&\times \frac{|I'|^K}{(|I'| + |u - x_{I'}|)^{(1+K)}} \frac{|I'|^K}{(|I'| + |v - y_{J'}|)^{(1+K)}} |I'||J'| |\phi_{R'} * f(x_{I'}, y_{J'})|^2,
\end{aligned}
$$

where K, L are any positive integers which can be chosen such that L, $K > 2/p - 1$ (for general $\mathbb{H}^n$, K can be chosen greater than $(2n + 2)(2/p - 1)$), the constant C depends only on K, L, and the functions ψ and ϕ, where $x_{I'}$ and $y_{J'}$, are any fixed points in I' and J', respectively.

Adding up over $R \subseteq \Omega$, we obtain

$$\sum_{R \subseteq \Omega} |I||J|S_R \leq C \sum_{R \subseteq \Omega} \sum_{R'} |I'||J'|r(R, R')P(R, R')T_{R'}, \tag{7-1}$$

where

$$r(R, R') = \left(\frac{|I|}{|I'|} \wedge \frac{|I'|}{|I|} \right)^{L-1} \left(\frac{|J|}{|J'|} \wedge \frac{|J'|}{|J|} \right)^{L-1}$$

and

$$P(R, R') = \frac{1}{(1 + \mathrm{dist}(I, I')/|I'|)^{1+K}(1 + \mathrm{dist}(J, J')/|J'|)^{1+K}}$$

if $j' > k'$, and

$$P(R, R') = \frac{1}{(1 + \mathrm{dist}(I, I')/|I'|)^{1+K}(1 + \mathrm{dist}(J, J')/|I'|)^{1+K}}$$

if $j' \leq k'$.

We estimate the right-hand side in the above inequality, where we first consider

$$R' = I' \times J', \quad |I'| = 2^{-j'-N}, \quad |J'| = 2^{-j'-N} + 2^{-k'-N}, \quad j' > k'.$$

Define

$$\Omega^{i,\ell} = \bigcup_{I \times J \subset \Omega} 3(2^i I \times 2^\ell J) \quad \text{for } i, \ell \geq 0.$$

Let $B_{i,\ell}$ be a collection of dyadic rectangles R' so that, for $i, \ell \geq 1$,

$$B_{i,\ell} = \left\{ R' = I' \times J' : 3(2^i I' \times 2^\ell J') \cap \Omega^{i,\ell} \neq \varnothing \text{ and } 3(2^{i-1} I' \times 2^{\ell-1} J') \cap \Omega^{i,\ell} = \varnothing \right\},$$

$$B_{0,\ell} = \left\{ R' = I' \times J' : 3(I' \times 2^\ell J') \cap \Omega^{0,\ell} \neq \varnothing \text{ and } 3(I' \times 2^{\ell-1} J') \cap \Omega^{0,\ell} = \varnothing \right\} \quad \text{for } \ell \geq 1,$$

$$B_{i,0} = \left\{ R' = I' \times J' : 3(2^i I' \times J') \cap \Omega^{i,0} \neq \varnothing \text{ and } 3(2^{i-1} I' \times J') \cap \Omega^{i,0} = \varnothing \right\} \quad \text{for } i \geq 1,$$

$$B_{0,0} = \left\{ R' = I' \times J' : 3(I' \times J') \cap \Omega^{0,0} \neq \varnothing \right\}.$$

We write

$$\sum_{R \subseteq \Omega} \sum_{R'} |I'||J'|r(R, R')P(R, R')T_{R'} = \sum_{\substack{i \geq 0 \\ \ell \geq 0}} \sum_{R' \in B_{i,\ell}} \sum_{R \subseteq \Omega} |I'||J'|r(R, R')P(R, R')T_{R'}.$$

To estimate the right-hand side of the above equality, we first consider the case when $i = \ell = 0$. Note that when $R' \in B_{0,0}$, $3R' \cap \Omega^{0,0} \neq \varnothing$. For each integer $h \geq 1$, let

$$\mathscr{F}_h = \left\{ R' = I' \times J' \in B_{0,0} : |(3I' \times 3J') \cap \Omega^{0,0}| \geq \frac{1}{2^h}|3I' \times 3J'| \right\}.$$

Let $\mathscr{D}_h = \mathscr{F}_h \backslash \mathscr{F}_{h-1}$, and $\Omega_h = \bigcup_{R' \in \mathscr{D}_h} R'$. Finally, assume that, for any open set $\Omega \subset R^2$,

$$\sum_{R = I \times J \subseteq \Omega} |I||J|T_R \le C|\Omega|^{2/p-1}.$$

Since $B_{0,0} = \bigcup_{h \ge 1} \mathscr{D}_h$ and for each $R' \in B_{0,0}$, $P(R, R') \le 1$, we have

$$\sum_{R' \in B_{0,0}} \sum_{R \subseteq \Omega} |I'||J'|r(R, R')P(R, R')T_{R'} \le \sum_{h \ge 1} \sum_{R' \subseteq \Omega_h} \sum_{R \subseteq \Omega} |I'||J'|r(R, R')T_{R'}.$$

For each $h \ge 1$ and $R' \subseteq \Omega_h$, we decompose $\{R : R \subseteq \Omega\}$ into

$$A_{0,0}(R') = \big\{ R = I \times J \subseteq \Omega : \operatorname{dist}(I, I') \le |I| \vee |I'|,\ \operatorname{dist}(J, J') \le |J| \vee |J'| \big\},$$

$$A_{i',0}(R') = \big\{ R = I \times J \subseteq \Omega : 2^{i'-1}(|I| \vee |I'|) < \operatorname{dist}(I, I') \le 2^{i'}(|I| \vee |I'|),\ \operatorname{dist}(J, J') \le |J| \vee |J'| \big\},$$

$$A_{0,\ell'}(R') = \big\{ R = I \times J \subseteq \Omega : \operatorname{dist}(I, I') \le |I| \vee |I'|,\ 2^{\ell'-1}(|J| \vee |J'|) < \operatorname{dist}(J, J') \le 2^{\ell'}(|J| \vee |J'|) \big\},$$

$$A_{i',\ell'}(R') = \big\{ R = I \times J \subseteq \Omega : 2^{i'-1}(|I| \vee |I'|) < \operatorname{dist}(I, I') \le 2^{i'}(|I| \vee |I'|),$$
$$2^{\ell'-1}(|J| \vee |J'|) < \operatorname{dist}(J, J') \le 2^{\ell'}(|J| \vee |J'|) \big\},$$

where $i', \ell' \ge 1$.

Now we split $\sum_{h \ge 1} \sum_{R' \subseteq \Omega_h} \sum_{R \subseteq \Omega} |I'||J'|r(R, R')P(R, R')T_{R'}$ into

$$\sum_{h \ge 1} \sum_{R' \in \Omega_h} \left(\sum_{R \in A_{0,0}(R')} + \sum_{i' \ge 1} \sum_{R \in A_{i',0}(R')} + \sum_{\ell' \ge 1} \sum_{R \in A_{0,\ell'}(R')} + \sum_{i',\ell' \ge 1} \sum_{R \in A_{i',\ell'}(R')} \right) |I'||J'|r(R, R')P(R, R')T_{R'}$$
$$=: I_1 + I_2 + I_3 + I_4.$$

To estimate the term I_1, we only need to estimate $\sum_{R \in A_{0,0}(R')} r(R, R')$, since $P(R, R') \le 1$ in this case.

Note that $R \in A_{0,0}(R')$ implies $3R \cap 3R' \ne \varnothing$. For such R, there are four cases:

Case 1: $|I'| \ge |I|$, $|J'| \le |J|$.

Case 2: $|I'| \le |I|$, $|J'| \ge |J|$.

Case 3: $|I'| \ge |I|$, $|J'| \ge |J|$.

Case 4: $|I'| \le |I|$, $|J'| \le |J|$.

In each case, we can estimate $\sum_{R \in A_{0,0}} r(R, R') \le C2^{-hL}$ by using a simple geometric argument similar to that in [Chang and Fefferman 1980]. This implies that I_1 is bounded by

$$\sum_{h \ge 1} 2^{-hL} |\Omega_h|^{2/p-1} \le C \sum_{h \ge 1} h^{2/p-1} 2^{-h(L-2/p+1)} |\Omega^{0,0}|^{2/p-1} \le C|\Omega|^{2/p-1},$$

since $|\Omega_h| \le Ch2^h |\Omega^{0,0}|$ and $|\Omega^{0,0}| \le C|\Omega|$.

Thus it remains to estimate term I_4, since estimates of I_2 and I_3 can be derived using the same techniques as for I_1 and I_4. The estimate for this term is more complicated than that for term I_1.

As in estimating term I_1, we only need to estimate the sum $\sum_{R \in A_{i',\ell'}(R')} r(R, R')$, since $P(R, R') \le 2^{-i(1+K)} 2^{-i'(1+K)}$. Note that $R \in A_{i',\ell'}(R')$ implies $3(2^{i'} I \times 2^{\ell'} J) \cap 3(2^{i'} I' \times 2^{\ell'} J') \ne \varnothing$. We also split our estimate into four cases.

Case 1. In this case, $|2^i I'| \geq |2^{i'} I|$, $|2^{\ell'} J'| \leq |2^{\ell'} J|$. Then

$$\frac{|2^{i'} I|}{|3 \cdot 2^{i'} I'|} |3(2^{i'} I' \times 2^{\ell'} J')| \leq |3(2^{i'} I' \times 2^{\ell'} J')| \wedge |3(2^i I \times 2^{\ell'} J)|$$

$$\leq C 2^{i'} 2^{\ell'} |3R' \cap \Omega^{0,0}| \leq C 2^{i'} 2^{\ell'} \frac{1}{2^{h-1}} |3R'| \leq C \frac{1}{2^{h-1}} |3(2^{i'} I' \times 2^{\ell'} J')|.$$

Thus $|2^{i'} I'| = \sum^{h-1+n} |2^i I|$ for some $n \geq 0$. For each fixed n, the number of such $2^{i'} I$ must be $\leq 2^n \cdot 5$. As for $|2^{\ell'} J| = 2^m |2^{\ell'} J'|$, for some $m \geq 0$, for each fixed m, $3 \cdot 2^{\ell'} J \cap 3 \cdot 2^{\ell'} J' \neq \varnothing$ implies that the number of such $2^{\ell'} J'$ is less than 5. Thus

$$\sum_{R \in \text{Case } 1} r(R, R') \leq \sum_{m,n \geq 0} \left(\frac{1}{2^{n+m+h-1}}\right)^L 2^n \cdot 5^2 \leq C 2^{-hL}.$$

We can handle the other three cases similarly. Combining the four cases, we have

$$\sum_{R \in A_{i',\ell'}(R')} r(R, R') \leq C 2^{-hL},$$

which, together with the estimate for $P(R, R')$, imply that

$$I_4 \leq C \sum_{h \geq 1} \sum_{i',\ell' \geq 1} \sum_{R' \subseteq \Omega_h} 2^{-hL} 2^{-i'(1+K)} 2^{-\ell'(1+K)} |I'||J'| T_{R'}.$$

Hence I_4 is bounded by

$$\sum_{h \geq 1} 2^{-hL} |\Omega_h|^{2/p-1} \leq C \sum_{h \geq 1} h^{2/p-1} 2^{-h(L-2/p+1)} |\Omega^{0,0}|^{2/p-1} \leq C |\Omega|^{2/p-1},$$

since $\sum_{R' \subset \Omega_h} |I'||J'| T_{R'} \leq C |\Omega_h|^{2/p-1}$ and $|\Omega_h| \leq C h 2^h |\Omega^{0,0}|$ and $|\Omega^{0,0}| \leq C |\Omega|$. Combining I_1, I_2, I_3, and I_4, we have

$$\frac{1}{|\Omega|^{2/p-1}} \sum_{R' \in B_{0,0}} \sum_{R \subseteq \Omega} |I'||J'| r(R, R') P(R, R') T_{R'} \leq C \sup_{\bar{\Omega}} 1/|\bar{\Omega}|^{2/p-1} \sum_{R' \subseteq \bar{\Omega}} |I'||J'| T_{R'}.$$

Now we consider

$$\sum_{i,\ell \geq 1} \sum_{R' \in B_{i,\ell}} \sum_{R \subseteq \Omega} |I'||J'| r(R, R') P(R, R') T_{R'}.$$

Note that for $R' \in B_{i,\ell}$, $3(2^i I' \times 2^\ell J) \cap \Omega_{i,\ell} \neq \varnothing$. Let

$$\mathscr{F}_h^{i,\ell} = \left\{ R' \in B_{i,\ell} : |3(2^i I' \times 2^\ell J') \cap \Omega^{i,\ell}| \geq \frac{1}{2^h} |3(2^i I' \times 2^\ell J')| \right\},$$

$$\mathscr{D}_h^{i,\ell} = \mathscr{F}_h^{i,\ell} \setminus \mathscr{F}_{h-1}^{i,\ell},$$

and

$$\Omega_h^{i,\ell} = \bigcup_{R' \in \mathscr{D}_h^{i,\ell}} R'.$$

Since $B_{i,\ell} = \bigcup_{h \geq 1} \mathscr{D}_h^{i,\ell}$, we first estimate

$$\sum_{R' \in \mathscr{D}_h^{i,\ell}} \sum_{R \subseteq \Omega} |I'||J'|r(R,R')P(R,R')T_{R'}$$

for some $i, \ell, h \geq 1$.

Note that for each $R' \in \mathscr{D}_h^{i,\ell}$, $3(2^i I \times 2^\ell J') \cap \Omega^{i-1,\ell-1} = \varnothing$. So, for any $R \subseteq \Omega$, we have $2^i(|I| \vee |I'|) \leq \mathrm{dist}(I, I')$ and $2^\ell(|J| \vee |J'|) \leq \mathrm{dist}(J, J')$. We decompose $\{R : R \subseteq \Omega\}$ as

$$A_{i',\ell'}(R') = \Big\{ R \subseteq \Omega : 2^{i'-1}2^i(|I| \vee |I'|) \leq \mathrm{dist}(I, I') \leq 2^{i'}2^i(|I| \vee |I'|),$$
$$2^{\ell'-1}2^\ell(|J| \vee |J'|) \leq \mathrm{dist}(J, J') \leq 2^{\ell'}2^\ell(|J| \vee |J'|) \Big\},$$

where $i', \ell' \geq 1$. Then we write

$$\sum_{R' \in \mathscr{D}_h^{i,\ell}} \sum_{R \subseteq \Omega} |I'||J'|r(R,R')P(R,R')T_{R'} = \sum_{i',\ell' \geq 1} \sum_{R' \in \mathscr{D}_h^{i,\ell}} \sum_{R \in A_{i',\ell'}(R')} |I'||J'|r(R,R')P(R,R')T_{R'}.$$

Since

$$P(R, R') \leq C 2^{-i(1+K)} 2^{-\ell(1+K)} 2^{-i'(1+K)} 2^{-\ell'(1+K)}$$

for $R' \in B_{i,\ell}$ and $R \in A_{i',\ell'}(R')$, repeating the same proof with $B_{0,0}$ replaced by $B_{i,\ell}$ and using the fact that

$$|\Omega_h^{i,\ell}| \leq C 2^h |\Omega^{i,\ell}|, \quad |\Omega^{i,\ell}| \leq C(i2^i)(\ell2^\ell)|\Omega^{0,0}|, \quad |\Omega^{0,0}| \leq C|\Omega|,$$

yield

$$\sum_{R' \in \mathscr{D}_h^{i,\ell}} \sum_{R \in A_{i',\ell'}(R')} |I'||J'|r(R,R')P(R,R')T_{R'}$$

$$\leq C 2^{-i(1+K)} 2^{-\ell(1+K)} 2^{-i'(1+K)} 2^{-\ell'(1+K)} |\Omega_h^{i,\ell}|^{2/p-1} \left(\frac{1}{|\Omega_h^{i,\ell}|^{2/p-1}} \sum_{R' \subset \Omega_h^{i,\ell}} |I'||J'|T_{R'} \right)$$

$$\leq C 2^{-i(1+K)} 2^{-\ell(1+K)} 2^{-i'(1+K)} 2^{-\ell'(1+K)} i^{2/p-1} 2^{i(2/p-1)} \ell^{2/p-1} 2^{\ell(2/p-1)} h^{2/p-1} 2^{-h(L-2/p+1)} |\Omega|^{2/p-1}$$

$$\times \sup_{\bar\Omega} \frac{1}{|\bar\Omega|^{2/p-1}} \sum_{R' \subset \bar\Omega} |I'||J'|T_{R'}.$$

Adding over all $i, \ell, i', \ell', h \geq 1$, we get

$$\frac{1}{|\Omega|^{2/p-1}} \sum_{i,\ell \geq 1} \sum_{R' \in B_{i,\ell}} \sum_{R \subseteq \Omega} |I'||J'|r(R,R')P(R,R')T_{R'} \leq C \sup_{\bar\Omega} \frac{1}{|\bar\Omega|^{2/p-1}} \sum_{R' \subseteq \bar\Omega} |I'||J'|T_{R'}.$$

Similar estimates, which we leave to the reader, hold for

$$\sum_{i \geq 1} \sum_{R' \in B_{i,0}} \sum_{R \subseteq \Omega} |I'||J'|r(R,R')P(R,R')T_{R'} \quad \text{and} \quad \sum_{\ell \geq 1} \sum_{R' \in B_{0,\ell}} \sum_{R \subseteq \Omega} |I'||J'|r(R,R')P(R,R')T_{R'},$$

which, after adding over all $i, \ell \geq 0$, completes the proof of Theorem 30. $\square$

As a consequence of Theorem 30, it is easy to see that the space CMO_F^p is well defined. In particular, we have the following:

Corollary 62. *We have*

$$\|f\|_{\mathrm{CMO}_F^p} \approx \sup_{\Omega} \left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_j \sum_k \sum_{I \times J \subseteq \Omega} |\psi_{j,k} * f(x_I, y_J)|^2 |I||J| \right\}^{\frac{1}{2}},$$

where $I \times J$ is a dyadic rectangle in $\mathbb{H}^n$ with $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-j-N} + 2^{-k-N}$, and where x_I, y_J are any fixed points in I, J, respectively.

Proof of Theorem 32. We first prove $c^p \subseteq (s^p)^*$. Applying the proof in Theorem 56, set

$$s(z, u) = \left\{ \sum_{I \times J} |s_{I \times J}|^2 |I|^{-1} |J|^{-1} \chi_I(z) \chi_J(u) \right\}^{\frac{1}{2}}$$

and

$$\Omega_i = \{(z, u) \in \mathbb{H}^n : s(z, u) > 2^i\}.$$

Let

$$\mathcal{B}_i = \left\{ (I \times J) : |(I \times J) \cap \Omega_i| > \tfrac{1}{2}|I \times J|, |(I \times J) \cap \Omega_{i+1}| \leq \tfrac{1}{2}|I \times J| \right\},$$

where $I \times J$ is a dyadic rectangle in $\mathbb{H}^n$ with $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-j-N} + 2^{-k-N}$. Suppose $t = \{t_{I \times J}\} \in c^p$, and write

$$\left| \sum_{I \times J} s_{I \times J} \bar{t}_{I \times J} \right| = \left| \sum_i \sum_{(I \times J) \in \mathcal{B}_i} s_{I \times J} \bar{t}_{I \times J} \right|$$

$$\leq \left\{ \sum_i \left[\sum_{(I \times J) \in \mathcal{B}_i} |s_{I \times J}|^2 \right]^{p/2} \left[\sum_{(I \times J) \in \mathcal{B}_i} |t_{I \times J}|^2 \right]^{p/2} \right\}^{1/p}$$

$$\leq C\|t\|_{c^p} \left\{ \sum_i |\Omega_i|^{1-p/2} \left[\sum_{(I \times J) \in \mathcal{B}_i} |s_{I \times J}|^2 \right]^{p/2} \right\}^{1/p}, \tag{7-2}$$

since if $I \times J \in \mathcal{B}_i$, then

$$I \times J \subseteq \widetilde{\Omega}_i = \left\{ (z, u) : M_S(\chi_{\Omega_i})(z, u) > \tfrac{1}{2} \right\},$$
$$|\widetilde{\Omega}_i| \leq C|\Omega_i|,$$

and $\{t_{I \times J}\} \in c^p$ yield

$$\left\{ \sum_{(I \times J) \in \mathcal{B}_i} |t_{I \times J}|^2 \right\}^{\frac{1}{2}} \leq C\|t\|_{c^p} |\Omega_i|^{1/p-1/2}.$$

The same proof as in the claim of Theorem 56 implies

$$\sum_{(I \times J) \in \mathcal{B}_i} |s_{I \times J}|^2 \leq C 2^{2i} |\Omega_i|.$$

Substituting the above term back into the last term in (7-2) gives $c^p \subseteq (s^p)^*$.

The proof of the converse is simple and is similar to the one given in [Frazier and Jawerth 1990] for $p = 1$ in the one-parameter setting on R^n. If $\ell \in (s^p)^*$, then it is clear that $\ell(s) = \sum_{I \times J} s_{I \times J} \bar{t}_{I \times J}$ for

some $t = \{t_{I \times J}\}$. Now fix an open set $\Omega \subset \mathbb{H}^n$ and let S be the sequence space of all $s = \{s_{I \times J}\}$ such that $I \times J \subseteq \Omega$. Finally, let μ be a measure on S so that the μ-measure of the "point" $I \times J$ is $1/|\Omega|^{2/p-1}$. Then,

$$
\left\{ \frac{1}{|\Omega|^{2/p-1}} \sum_{I \times J \subseteq \Omega} |t_{I \times J}|^2 \right\}^{\frac{1}{2}} = \|t_{I \times J}\|_{\ell^2(S, d\mu)}
$$

$$
= \sup_{\|s\|_{\ell^2(S, d\mu)} \leq 1} \left| \frac{1}{|\Omega|^{2/p-1}} \sum_{I \times J \subseteq \Omega} s_{I \times J} \bar{t}_{I \times J} \right|
$$

$$
\leq \|t\|_{(s^p)^*} \sup_{\|s\|_{\ell^2(S, d\mu)} \leq 1} \left\| s_{I \times J} \frac{1}{|\Omega|^{2/p-1}} \right\|_{s^p}.
$$

By Hölder's inequality,

$$
\left\| s_{I \times J} \frac{1}{|\Omega|^{2/p-1}} \right\|_{s^p} = \frac{1}{|\Omega|^{2/p-1}} \left\{ \int_\Omega \left(\sum_{I \times J \subseteq \Omega} |s_{I \times J}|^2 |I \times J|^{-1} \chi_I(x) \chi_J(y) \right)^{p/2} dz\, du \right\}^{1/p}
$$

$$
\leq \left\{ \frac{1}{|\Omega|^{2/p-1}} \int_\Omega \sum_{I \times J \subseteq \Omega} |s_{I \times J}|^2 |I \times J|^{-1} \chi_I(x) \chi_J(y)\, dz\, du \right\}^{1/2} = \|s\|_{\ell^2(S, d\mu)} \leq 1,
$$

which shows $\|t\|_{c^p} \leq \|t\|_{(s^p)^*}$. $\qquad \square$

In order to use Theorem 32 to obtain Theorem 33, we introduce a map S which takes $f \in (\mathcal{M}_{\text{flag}}^{M+\delta})'$ to the sequence of coefficients

$$
Sf \equiv \{s_{I \times J}\} = \left\{ |I|^{\frac{1}{2}} |J|^{\frac{1}{2}} \psi_{j,k} * f(x_I, y_J) \right\},
$$

where $I \times J$ is a dyadic rectangle in $\mathbb{H}^n$ with $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-j-N} + 2^{-k-N}$, and where x_I, y_J are any fixed points in I, J, respectively. For any sequence $s = \{s_{I \times J}\}$, we define a map T which takes s to

$$
T(s) = \sum_j \sum_k \sum_J \sum_I |I|^{\frac{1}{2}} |J|^{\frac{1}{2}} \tilde{\psi}_{j,k}(z, u) s_{I \times J},
$$

where the $\tilde{\psi}_{j,k}$ are as in (3-1).

The following result together with Theorem 32 will give Theorem 33.

Theorem 63. *The maps* $S : H_{\text{flag}}^p \to s^p$ *and* $S : \text{CMO}_{\text{flag}}^p \to c^p$, *as well as the maps* $T : s^p \to H_{\text{flag}}^p$ *and* $T : c^p \to \text{CMO}_{\text{flag}}^p$, *are bounded. Moreover,* $T \circ S$ *is the identity on both* H_{flag}^p *and* $\text{CMO}_{\text{flag}}^p$.

Proof. The boundedness of S on H_{flag}^p and $\text{CMO}_{\text{flag}}^p$ follows directly from the Plancherel–Pólya inequalities, Theorems 19 and 30. The boundedness of T also follows from the arguments in Theorems 19 and 30. Indeed, to see that T is bounded from s^p to H_{flag}^p, let $s = \{s_{I \times J}\}$. Then, by Proposition 54,

$$
\|T(s)\|_{H_{\text{flag}}^p} \leq C \left\| \left\{ \sum_j \sum_k \sum_J \sum_I |\psi_{j,k} * T(s)(z, u)|^2 \chi_I(x) \chi_J(y) \right\}^{\frac{1}{2}} \right\|_p.
$$

By adapting an argument similar to that in the proof of Theorem 19, we have, for some $0 < r < p$,

$$|\psi_{j,k} * T(s)(z, u)\chi_I(x)\chi_J(y)|^2$$

$$= \left|\sum_{j',k'}\sum_{I',J'}|I'||I'|\psi_{j,k}*\tilde{\psi}_{j',k'}(\cdot,\cdot)(z,u)s_{I'\times J'}|I'|^{-\frac{1}{2}}|J'|^{-\frac{1}{2}}\chi_I(x)\chi_J(y)\right|^2$$

$$\leq C\sum_{k\wedge k'\leq j\wedge j'}2^{-|j-j'|K}2^{-|k-k'|K}\left\{M_S\left(\sum_{I',J'}|s_{I'\times J'}||I'|^{-1}|J'|^{-1}\chi_{J'\chi_{I'}}\right)^r\right\}^{2/r}(z,u)\chi_I(x)\chi_J(y)$$

$$+ \sum_{k\wedge k'>j\wedge j'}2^{-|j-j'|K}2^{-|k-k'|K}\left\{M\left(\sum_{I',J'}|s_{I'\times J'}||I'|^{-1}|J'|^{-1}\chi_{J'\chi_{I'}}\right)^r\right\}^{2/r}(z,u)\chi_I(x)\chi_J(y).$$

Repeating the argument in Theorem 19 gives the boundedness of T from s^p to H^p_{flag}. A similar adaptation of the argument in the proof of Theorem 30 applies to yield the boundedness of T from c^p to $\text{CMO}^p_{\text{flag}}$. We leave the details to the reader. The discrete Calderón reproducing formula and Theorems 17 and 30 show that $T \circ S$ is the identity on both H^p_{flag} and $\text{CMO}^p_{\text{flag}}$. $\square$

We are now ready to give the proofs of Theorems 33 and 34.

Proof of Theorem 33. If $f \in \mathcal{M}^{M+\delta}_{\text{flag}}$ and $g \in \text{CMO}^p_{\text{flag}}$, let $\ell_g = \langle f, g\rangle$. Then the discrete Calderón reproducing formula and Theorems 30 and 32 imply

$$|\ell_g| = |\langle f, g\rangle| = \left|\sum_{R=I\times J}|I||J|\psi_R*f(x_I, y_J)\tilde{\psi}_R(g)(x_I, y_J)\right| \leq C\|f\|_{H^p_{\text{flag}}}\|g\|_{\text{CMO}^p_F}.$$

Because $\mathcal{M}^{M+\delta}_{\text{flag}}$ is dense in H^p_{flag}, this shows that the map $\ell_g = \langle f, g\rangle$, defined initially for $f \in \mathcal{M}^{M+\delta}_{\text{flag}}$, can be extended to a continuous linear functional on H^p_{flag} with $\|\ell_g\| \leq C\|g\|_{\text{CMO}^p_{\text{flag}}}$.

Conversely, let $\ell \in (H^p_{\text{flag}})^*$ and set $\ell_1 = \ell \circ T$, where T is defined as in Theorem 32. Then, by Theorem 32, $\ell_1 \in (s^p)^*$, so by Theorem 30, there exists $t = \{t_{I\times J}\}$ such that $\ell_1(s) = \sum_{I\times J}s_{I\times J}\bar{t}_{I\times J}$ for all $s = \{s_{I\times J}\}$, and where

$$\|t\|_{c^p} \approx \|\ell_1\| \leq C\|\ell\|,$$

because T is bounded. Again by Theorem 32, $\ell = \ell \circ T \circ S = \ell_1 \circ S$. Hence, with

$$f \in \mathcal{M}^{M+\delta}_{\text{flag}} \quad \text{and} \quad g = \sum_{I\times J}t_{I\times J}\psi_R((z, u)\circ(x_I, y_J)^{-1}),$$

and where without loss the generality we may assume that ψ is a radial function, we have

$$\ell(f) = \ell_1(S(f)) = \langle S(f), t\rangle = \langle f, g\rangle.$$

This proves $\ell = \ell_g$, and by Theorem 32 we have

$$\|g\|_{\text{CMO}^p_{\text{flag}}} \leq C\|t\|_{c^p} \leq C\|\ell_g\|.$$ $\square$

Proof of Theorem 34. As mentioned earlier, H^1_{flag} is a subspace of L^1. By the duality of H^1_{flag} and BMO_{flag}, we now conclude that L^∞ is a subspace of BMO_{flag}, and from the boundedness of flag singular integrals on H^1_{flag}, we get that flag singular integrals are bounded on BMO_{flag} and also from L^∞ to BMO_{flag}. $\square$

8. Calderón–Zygmund decomposition and interpolation decomposition

In this section we derive a Calderón–Zygmund decomposition using functions in flag Hardy spaces. As an application, we prove an interpolation theorem for the spaces $H^p_{\text{flag}}(\mathbb{H}^n)$.

We first recall that Chang and Fefferman [1982] established the following Calderón–Zygmund decomposition on the pure product domain $\mathbb{R}^2_+ \times \mathbb{R}^2_+$.

Lemma 64 (Calderón–Zygmund lemma). *Let $\alpha > 0$ be given and $f \in L^p(\mathbb{R}^2)$, $1 < p < 2$. Then we may write $f = g + b$, where $g \in L^2(\mathbb{R}^2)$ and $b \in H^1(\mathbb{R}^2_+ \times \mathbb{R}^2_+)$ with $\|g\|^2_2 \leq \alpha^{2-p}\|f\|^p_p$ and $\|b\|_{H^1(\mathbb{R}^2_+ \times \mathbb{R}^2_+)} \leq C\alpha^{1-p}\|f\|^p_p$, where c is an absolute constant.*

We now prove the Calderón–Zygmund decomposition in the setting of flag Hardy spaces on the Heisenberg group.

Proof of Theorem 35. We first assume $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$. Let $\alpha > 0$ and

$$\Omega_\ell = \{(z, u) \in \mathbb{H}^n : S(f)(z, u) > \alpha 2^\ell\},$$

where, as in Corollary 60,

$$S(f)(z, u) = \left\{\sum_{j,k} \sum_{I,J} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I(x)\chi_J(y)\right\}^{\frac{1}{2}}.$$

It was shown in Corollary 60 that for $f \in L^2(\mathbb{H}^n) \cap H^p_{\text{flag}}(\mathbb{H}^n)$, we have $\|f\|_{H^p_{\text{flag}}} \approx \|S(f)\|_p$.

In the following, we denote dyadic rectangles in $\mathbb{H}^n$ by $R = I \times J$ with $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-j-N} + 2^{-k-N}$, where j, k are integers and N is sufficiently large. Let

$$\mathcal{R}_0 = \left\{R = I \times J : |R \cap \Omega_0| < \tfrac{1}{2}|R|\right\}$$

and, for $\ell \geq 1$,

$$\mathcal{R}_\ell = \left\{R = I \times J : |R \cap \Omega_{\ell-1}| \geq \tfrac{1}{2}|R| \text{ but } |R \cap \Omega_\ell| < \tfrac{1}{2}|R|\right\}.$$

By the discrete Calderón reproducing formula in Theorem 57,

$$
\begin{aligned}
f(z, u) &= \sum_{j,k} \sum_{I,J} |I||J|\tilde{\phi}_{jk}((z, u) \circ (x_I, y_J)^{-1})\phi_{jk} * (T_N^{-1}(f))(x_I, y_J) \\
&= \sum_{\ell \geq 1} \sum_{I \times J \in \mathcal{R}_\ell} |I||J|\tilde{\phi}_{jk}((z, u) \circ (x_I, y_J)^{-1})\phi_{jk} * (T_N^{-1}(f))(x_I, y_J) \\
&\qquad\qquad + \sum_{I \times J \in \mathcal{R}_0} |I||J|\tilde{\phi}_{jk}((z, u) \circ (x_I, y_J)^{-1})\phi_{jk} * (T_N^{-1}(f))(x_I, y_J) \\
&= b(z, u) + g(z, u)
\end{aligned}
$$

When $p_1 > 1$, using a duality argument it is easy to show

$$\|g\|_{p_1} \leq C \left\|\left\{\sum_{R=I \times J \in \mathcal{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I \chi_J\right\}^{\frac{1}{2}}\right\|_{p_1}.$$

Next, we estimate $\|g\|_{H^{p_1}_{\text{flag}}}$ when $0 < p_1 \leq 1$. Clearly, the duality argument will not work here. Nevertheless, we can estimate the $H^{p_1}_{\text{flag}}$ norm directly by using the discrete Calderón reproducing formula in Theorem 57. To this end, we note that

$$
\|g\|_{H^{p_1}_{\text{flag}}} \leq \left\| \left\{ \sum_{j',k'} \sum_{I',J'} |(\psi_{j'k'} * g)(x_{I'}, y_{J'})|^2 \chi_{I'}(z) \chi_{J'}(u) \right\}^{\frac{1}{2}} \right\|_{L^{p_1}}.
$$

Since

$$
(\psi_{j',k'} * g)(x_{I'}, y_{J'}) = \sum_{I \times J \in \mathscr{R}_0} |I||J|(\psi_{j'k'} * \tilde{\phi}_{jk})((x_{I'}, y_{J'}) \circ (x_I, y_J)^{-1}) \phi_{jk} * (T_N^{-1}(f))(x_I, y_J),
$$

we can repeat the argument in the proof of Theorem 56 to obtain

$$
\left\| \left\{ \sum_{j',k'} \sum_{I',J'} |(\psi_{j'k'} * g)(x_{I'}, y_{J'})|^2 \chi_{I'}(z) \chi_{J'}(u) \right\}^{\frac{1}{2}} \right\|_{L^{p_1}} \leq C \left\| \left\{ \sum_{R = I \times J \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_{p_1}.
$$

This shows that for all $0 < p_1 < \infty$,

$$
\|g\|_{H^{p_1}_{\text{flag}}} \leq C \left\| \left\{ \sum_{R = I \times J \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_{p_1}.
$$

Claim 65. *We have*

$$
\int_{S(f)(z,u) \leq \alpha} S^{p_1}(f)(z, u)\, dz\, du \geq C \left\| \left\{ \sum_{R = I \times J \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_{p_1}.
$$

This claim implies

$$
\|g\|_{p_1} \leq C \int_{S(f)(z,u) \leq \alpha} S^{p_1}(f)(z, u)\, dz\, du \leq C\alpha^{p_1 - p} \int_{S(f)(z,u) \leq \alpha} S^p(f)(z, u)\, dz\, du \leq C\alpha^{p_1 - p} \|f\|^p_{H^p_{\text{flag}}(\mathbb{H}^n)}.
$$

To prove Claim 65, we let $R = I \times J \in \mathscr{R}_0$. Choose $0 < q < p_1$ and note that

$$
\int_{S(f)(z,u) \leq \alpha} S^{p_1}(f)(z, u)\, dz\, du
$$

$$
= \int_{S(f)(z,u) \leq \alpha} \left\{ \sum_{j,k} \sum_{I,J} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I(x) \chi_J(y) \right\}^{p_1/2} dz\, du
$$

$$
\geq C \int_{\Omega_0^c} \left\{ \sum_{R \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I \chi_J \right\}^{p_1/2} dz\, du
$$

$$
= C \int_{\mathbb{H}^n} \left\{ \sum_{R \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_{R \cap \Omega_0^c}(z, u) \right\}^{p_1/2} dz\, du
$$

$$
\geq C \int_{\mathbb{H}^n} \left\{ \left\{ \sum_{R \in \mathscr{R}_0} (M_S(|\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^q \chi_{R \cap \Omega_0^c})(z, u))^{2/q} \right\}^{q/2} \right\}^{p_1/q} dz\, du
$$

$$
\geq C \int_{\mathbb{H}^n} \left\{ \sum_{R \in \mathscr{R}_0} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_R(z, u) \right\}^{p_1/2} dz\, du.
$$

In the last inequality above we have used the fact that $|\Omega_0^c \cap (I \times J)| \geq \frac{1}{2}|I \times J|$ for $I \times J \in \mathcal{R}_0$, and thus

$$\chi_R(z, u) \leq 2^{1/q} M_S(\chi_{R \cap \Omega_0^c})^{1/q}(z, u).$$

In the second-to-last inequality above we have used the vector-valued Fefferman–Stein inequality for the strong maximal function

$$\left\| \left(\sum_{k=1}^{\infty} (M_S f_k)^r \right)^{1/r} \right\|_p \leq C \left\| \left(\sum_{k=1}^{\infty} |f_k|^r \right)^{1/r} \right\|_p,$$

with the exponents $r = 2/q > 1$ and $p = p_1/q > 1$. Thus Claim 65 follows.

We now recall that $\widetilde{\Omega}_\ell = \left\{ (z, u) \in \mathbb{H}^n : M_S(\chi_{\Omega_\ell}) > \frac{1}{2} \right\}$.

Claim 66. *For $p_2 \leq 1$,*

$$\left\| \sum_{I \times J \in \mathcal{R}_\ell} |I||J| \widetilde{\phi}_{jk}((x, y) \circ (x_I, y_J)^{-1}) \phi_{jk} * (T_N^{-1} f)(x_I, y_J) \right\|_{H_{\text{flag}}^{p_2}}^{p_2} \leq C(2^\ell \alpha)^{p_2} |\widetilde{\Omega_{\ell-1}}|.$$

Claim 66 implies

$$\begin{aligned}
\|b\|_{H_{\text{flag}}^{p_2}}^{p_2} &\leq \sum_{\ell \geq 1} (2^\ell \alpha)^{p_2} |\widetilde{\Omega_{\ell-1}}| \\
&\leq C \sum_{\ell \geq 1} (2^\ell \alpha)^{p_2} |\Omega_{\ell-1}| \\
&\leq C \int_{S(f)(z,u) > \alpha} S^{p_2} f(z, u) \, dz \, du \\
&\leq C \alpha^{p_2 - p} \int_{S(f)(z,u) > \alpha} S^p f(z, u) \, dz \, du \leq C \alpha^{p_2 - p} \|f\|_{H_{\text{flag}}^p}^p.
\end{aligned}$$

To prove Claim 66, we again have

$$\begin{aligned}
&\left\| \sum_{I \times J \in \mathcal{R}_\ell} |I||J| \widetilde{\phi}_{jk}((x, y) \circ (x_I, y_J)^{-1}) \phi_{jk} * (T_N^{-1} f)(x_I, y_J) \right\|_{H_{\text{flag}}^{p_2}}^{p_2} \\
&\leq C \left\| \left\{ \sum_{j'k'} \sum_{I', J'} \left| \sum_{I \times J \in \mathcal{R}_\ell} |I||J| (\psi_{j'k'} * \widetilde{\phi}_{jk})((x_{I'}, y_{J'}) \circ (x_I, y_J)^{-1}) \phi_{jk} * (T_N^{-1} f)(x_I, y_J) \right|^2 \right\}^{\frac{1}{2}} \right\|_{L^{p_2}} \\
&\leq C \left\| \left\{ \sum_{R = I \times J \in \mathcal{R}_\ell} |\phi_{jk} * (T_N^{-1} f)(x_I, y_J)|^2 \chi_I \chi_J \right\}^{\frac{1}{2}} \right\|_{p_2},
\end{aligned}$$

where we can use an argument similar to that in the proof of Theorem 56 to prove the last inequality.

However,

$$\sum_{\ell=1}^{\infty} (2^{\ell}\alpha)^{p_2} |\widetilde{\Omega}_{\ell-1}| \geq \int_{\widetilde{\Omega}_{\ell-1}\setminus\Omega_{\ell}} S(f)^{p_2}(z, u) \, dz \, du$$

$$= \int_{\widetilde{\Omega}_{\ell-1}\setminus\Omega_{\ell}} \left\{ \sum_{j,k} \sum_{I,J} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I(z)\chi_J(u) \right\}^{p_2/2} dz \, du$$

$$= \int_{\mathbb{H}^n} \left\{ \sum_{j,k} \sum_{I,J} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_{(I\times J)\cap\widetilde{\Omega}_{\ell-1}\setminus\Omega_{\ell}}(z, u) \right\}^{p_2/2} dz \, du$$

$$\geq \int_{\mathbb{H}^n} \left\{ \sum_{I\times J\in\mathscr{R}_{\ell}} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_{(I\times J)\cap\widetilde{\Omega}_{\ell-1}\setminus\Omega_{\ell}}(z, u) \right\}^{p_2/2} dz \, du$$

$$\geq \int_{\mathbb{H}^n} \left\{ \sum_{I\times J\in\mathscr{R}_{\ell}} |\phi_{jk} * (T_N^{-1}(f))(x_I, y_J)|^2 \chi_I(z)\chi_J(u) \right\}^{p_2/2} dz \, du.$$

In the above string of inequalities, we have used the fact that, for $R \in \mathscr{R}_{\ell}$, we have

$$|R \cap \Omega_{\ell-1}| > \tfrac{1}{2}|R| \quad \text{and} \quad |R \cap \Omega_{\ell}| \leq \tfrac{1}{2}|R|,$$

and, consequently, $R \subset \widetilde{\Omega}_{\ell-1}$. Therefore

$$|R \cap (\widetilde{\Omega}_{\ell-1}\setminus\Omega_{\ell})| > \tfrac{1}{2}|R|.$$

Thus the same argument applies here to conclude the last inequality above. Finally, since

$$L^2(\mathbb{H}^n) \cap H_{\text{flag}}^p(\mathbb{H}^n)$$

is dense in $H_{\text{flag}}^p(\mathbb{H}^n)$, Theorem 35 is proved. $\qquad\square$

We are now ready to prove the interpolation theorem on Hardy spaces H_{flag}^p for all $0 < p < \infty$.

Proof of Theorem 36. Suppose that T is bounded from $H_{\text{flag}}^{p_2}$ to L^{p_2} and from $H_{\text{flag}}^{p_1}$ to L^{p_1}. For any given $\lambda > 0$ and $f \in H_{\text{flag}}^p$, by the Calderón–Zygmund decomposition,

$$f(z, u) = g(z, u) + b(z, u)$$

with

$$\|g\|_{H_{\text{flag}}^{p_1}}^{p_1} \leq C\lambda^{p_1-p} \|f\|_{H_{\text{flag}}^p}^p \quad \text{and} \quad \|b\|_{H_{\text{flag}}^{p_2}}^{p_2} \leq C\lambda^{p_2-p} \|f\|_{H_{\text{flag}}^p}^p.$$

Moreover, we have proved the estimates

$$\|g\|_{H_{\text{flag}}^{p_1}}^{p_1} \leq C \int_{S(f)(z,u)\leq\alpha} S(f)^{p_1}(z, u) \, dz \, du \quad \text{and} \quad \|b\|_{H_{\text{flag}}^{p_2}}^{p_2} \leq C \int_{S(f)(z,u)>\alpha} S(f)^{p_2}(z, u) \, dz \, du,$$

which imply that

$$\|Tf\|_p^p = p \int_0^\infty \alpha^{p-1} \big| \{(z,u) : |Tf(z,u)| > \lambda\} \big| \, d\alpha$$

$$\leq p \int_0^\infty \alpha^{p-1} \big| \{(z,u) : |Tg(z,u)| > \tfrac{1}{2}\lambda\} \big| \, d\alpha + p \int_0^\infty \alpha^{p-1} \big| \{(z,u) : |Tb(z,u)| > \tfrac{1}{2}\lambda\} \big| \, d\alpha$$

$$\leq p \int_0^\infty \alpha^{p-1} \int_{S(f)(z,u)\leq\alpha} S(f)^{p_1}(z,u) \, dz \, du \, d\alpha + p \int_0^\infty \alpha^{p-1} \int_{S(f)(z,u)>\alpha} S(f)^{p_2}(z,u) \, dz \, du \, d\alpha$$

$$\leq C \|f\|_{H_{\text{flag}}^p}^p.$$

Thus,

$$\|Tf\|_p \leq C \|f\|_{H_{\text{flag}}^p}$$

for any $p_2 < p < p_1$. Hence T is bounded from H_{flag}^p to L^p.

Now we prove the second assertion, that T is bounded on H_{flag}^p for $p_2 < p < p_1$. For any given $\lambda > 0$ and $f \in H_{\text{flag}}^p$, we have, again by the Calderón–Zygmund decomposition,

$$\big| \{(z,u) : |g(Tf)(z,u)| > \alpha\} \big|$$

$$\leq \big| \{(z,u) : |g(Tg)(z,u)| > \tfrac{1}{2}\alpha\} \big| + \big| \{(z,u) : |g(Tb)(z,u)| > \tfrac{1}{2}\alpha\} \big|$$

$$\leq C\alpha^{-p_1} \|Tg\|_{H_{\text{flag}}^{p_1}}^{p_1} + C\alpha^{-p_2} \|Tb\|_{H_{\text{flag}}^{p_2}}^{p_2}$$

$$\leq C\alpha^{-p_1} \|g\|_{H_{\text{flag}}^{p_1}}^{p_1} + C\alpha^{-p_2} \|b\|_{H_{\text{flag}}^{p_2}}^{p_2}$$

$$\leq C\alpha^{-p_1} \int_{S(f)(z,u)\leq\alpha} (Sf)^{p_1}(z,u) \, dz \, du + C\alpha^{-p_2} \int_{S(f)(z,u)>\alpha} (Sf)^{p_2}(z,u) \, dz \, du,$$

which, as above, shows that $\|Tf\|_{H_{\text{flag}}^p} \leq C \|g(TF)\|_p \leq C \|f\|_{H_{\text{flag}}^p}$ for any $p_2 < p < p_1$. $\qquad\square$

9. A counterexample for the one-parameter Hardy space

Recall that $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$ is the Heisenberg group with group multiplication

$$(\zeta, t) \cdot (\eta, s) = (\zeta + \eta, t + s + 2\operatorname{Im}(\zeta \cdot \bar{\eta})), \quad (\zeta, t), (\eta, s) \in \mathbb{C}^n \times \mathbb{R},$$

and that $(\eta, s)^{-1} = (-\eta, -s)$. Consider the mixed kernel $K(z,t) = K_1(z)K_2(z,t)$ for $(z,t) \in \mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$ given by

$$K_1(z) = \frac{\Omega(z)}{|z|^{2n}} \quad \text{and} \quad K_2(z,t) = \frac{1}{|z|^2 + it},$$

where Ω is smooth with mean zero on the unit sphere in $\mathbb{C}^n$. We show in the subsection below that K satisfies the smoothness and cancellation conditions required of a flag kernel. It then follows from [Müller et al. 1995] that there is an operator T having kernel K such that, for each $1 < p < \infty$,

$$\|Tf\|_{L^p(\mathbb{H}^n)} \leq C_{p,n} \|f\|_{L^p(\mathbb{H}^n)}, \quad f \in L^p(\mathbb{H}^n).$$

The action of the corresponding singular integral operator $Tf = K * f$ is given by

$$Tf(\zeta, t) = K *_{\mathbb{H}^n} f(\zeta, t) = \int_{\mathbb{H}^n} K((\zeta, t) \circ (\eta, s)^{-1}) f(\eta, s) \, d\eta \, ds$$

$$= \int_{\mathbb{H}^n} f(\eta, s) K(\zeta - \eta, t - s - 2 \operatorname{Im}(\zeta \cdot \bar{\eta})) \, d\eta \, ds$$

$$= \int_{\mathbb{H}^n} f(\eta, s) \frac{\Omega(\zeta - \eta)}{|\zeta - \eta|^{2n}} \frac{1}{|\zeta - \eta|^2 + i(t - s - 2 \operatorname{Im}(\zeta \cdot \bar{\eta}))} \, d\eta \, ds.$$

Theorem 67. *There is a smooth function Ω with mean zero on the unit sphere in $\mathbb{C}^n$ such that there is no operator T having kernel K that is bounded from $H^1(\mathbb{H}^n)$ to $L^1(\mathbb{H}^n)$.*

To prove the theorem, we fix $f(z, u) = \psi(z)\varphi(u)$, where

(1) ψ is smooth with support in the unit ball of $\mathbb{C}^n$,

(2) φ is smooth with support in $(-1, 1)$,

(3) $\int_{\mathbb{C}^n} \psi(z) \, dz = 0$ and $\int_{\mathbb{R}} \varphi(u) \, du = 1$.

Such a function f is clearly in $H^1(\mathbb{H}^n)$ since f is smooth, compactly supported, and has mean zero:

$$\int_{\mathbb{H}^n} f(z, u) \, dz \, du = \int_{\mathbb{R}} \left\{ \int_{\mathbb{C}^n} \psi(z) \, dz \right\} \varphi(u) \, du = \int_{\mathbb{R}} \{0\} \varphi(u) \, du = 0.$$

We next show that T fails to be bounded from $H^1(\mathbb{H}^n)$ to $L^1(\mathbb{H}^n)$, and then that T is a flag singular integral.

9.1. *Failure of boundedness of T.* For

$$\zeta \in B((100, \mathbf{0}), 0) = \{(\zeta_1, \zeta') \in \mathbb{R} \times \mathbb{C}^{n-1} : (\zeta_1 - 100)^2 + |\zeta'|^2 < 1\}, \quad |t| > 10^6,$$

we have

$$|Tf(\zeta, t)| \approx \int \psi(\eta)\varphi(s) \frac{\Omega(\zeta - \eta)}{|\zeta|^{2n}} \frac{1}{|\zeta|^2 + i(t - 2|\zeta|^2)} \, d\eta \, ds \approx \frac{1}{|\zeta|^{2n}|t|},$$

since, for $\zeta \in B((100, \mathbf{0}), 0)$, we have

$$\left| \int \psi(\eta)\Omega(\zeta - \eta) \, d\eta \right| \geq c > 0,$$

for an appropriately chosen Ω with mean zero on the sphere. The point is that both functions ψ and Ω have mean zero on their respective domains, but the product can destroy enough of the cancellation. For example, when $n = 1$, we can take

$$\Omega(x, y) = \frac{y}{\sqrt{x^2 + y^2}},$$

$$\psi(x, y) = y\psi_1(x)\psi_2(y),$$

where ψ_i is an even function identically one on $\left(-1/2, 1/2\right)$ and supported in $(-1/\sqrt{2}, 1/\sqrt{2})$. Then, for

$$\zeta = (100 + v, \omega), \quad |v|^2 + |\omega|^2 \leq 1,$$

we have

$$\int \psi(\eta)\Omega(\zeta - \eta)\, d\eta = \int y\psi_1(x)\psi_2(y)\Omega(100 + \nu - x, \omega - y)$$

$$= \int y\psi_1(x)\psi_2(y)\frac{\omega - y}{\sqrt{(100 + \nu - x)^2 + (\omega - y)^2}}$$

$$= \omega\int \frac{y\psi_1(x)\psi_2(y)}{\sqrt{(100 + \nu - x)^2 + (\omega - y)^2}} - \int \frac{y^2\psi_1(x)\psi_2(y)}{\sqrt{(100 + \nu - x)^2 + (\omega - y)^2}}$$

$$\approx -\frac{1}{100}.$$

We conclude from the above that

$$\int_{\mathbb{H}^n} |Tf(\zeta, t)|\, d\zeta\, dt \gtrsim \int_{\{\zeta \in B((100,\mathbf{0}),0) \text{ and } |t| > 10^6\}} \frac{1}{|\zeta|^{2n}|t|}\, d\zeta\, dt = \infty.$$

9.2. T is a flag singular integral. Let K be the kernel

$$K(z, t) = \frac{\Omega(z)}{|z|^{2n}}\frac{1}{|z|^2 + it}, \quad (z, t) \in \mathbb{H}^n.$$

In order to show that K is a flag kernel, we must establish the following smoothness and cancellation conditions.

(1) (differential inequalities) For any multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\beta = (\beta_1, \ldots, \beta_m)$,

$$|\partial_z^\alpha \partial_u^\beta K(z, u)| \leq C_{\alpha, \beta}|z|^{-2n - |\alpha|} \cdot (|z|^2 + |u|)^{-1 - |\beta|}$$

for all $(z, u) \in \mathbb{H}^n$ with $z \neq 0$.

(2) (cancellation condition) For every multi-index α, every normalized bump function ϕ_1 on $\mathbb{R}$, and every $\delta > 0$,

$$\left|\int_{\mathbb{R}} \partial_z^\alpha K(z, u)\phi_1(\delta u)\, du\right| \leq C_\alpha|z|^{-2n - |\alpha|};$$

for every multi-index β, every normalized bump function ϕ_2 on $\mathbb{C}^n$, and every $\delta > 0$,

$$\left|\int_{\mathbb{C}^n} \partial_u^\beta K(z, u)\phi_2(\delta z)\, dz\right| \leq C_\gamma|u|^{-1 - |\beta|};$$

and for every normalized bump function ϕ_3 on $\mathbb{H}^n$ and every $\delta_1 > 0$ and $\delta_2 > 0$,

$$\left|\int_{\mathbb{H}^n} K(z, u)\phi_3(\delta_1 z, \delta_2 u)\, dz\, du\right| \leq C.$$

The differential inequalities in (1) follow immediately from the definition of K.

The first cancellation condition in (2) exploits the fact that t is an odd function. For convenience we assume $\alpha = 0$. We then have

$$
\left| \int_{\mathbb{R}} K(z,t)\phi_1(\delta t)\, dt \right| = \left| \int_{\mathbb{R}} \frac{\Omega(z)}{|z|^{2n}} \left\{ \frac{|z|^2}{|z|^4 + t^2} - \frac{it}{|z|^4 + t^2} \right\} \phi_1(\delta t)\, dt \right|
$$

$$
\leq \int_{\mathbb{R}} \frac{1}{|z|^{2n}} \frac{|z|^2}{|z|^4 + t^2} |\phi_1(\delta t)|\, dt + \left| \int_{\mathbb{R}} \frac{\Omega(z)}{|z|^{2n}} \frac{it}{|z|^4 + t^2} \{\phi_1(\delta t) - \phi_1(0)\}\, dt \right|
$$

$$
\lesssim \frac{1}{|z|^{2n-2}} \int_0^\infty \frac{1}{|z|^4 + t^2}\, dt + \frac{1}{|z|^{2n}} \int_0^{1/\delta} \frac{\delta t^2}{|z|^4 + t^2}\, dt.
$$

Now

$$
\frac{1}{|z|^{2n-2}} \int_0^\infty \frac{1}{|z|^4 + t^2}\, dt \lesssim \frac{1}{|z|^{2n-2}} \left(\int_0^{|z|^2} \frac{1}{|z|^4}\, dt + \int_{|z|^2}^\infty \frac{1}{t^2}\, dt \right) \lesssim \frac{1}{|z|^{2n}},
$$

and, for $|z|^2 \leq 1/\delta$, we have

$$
\int_0^{1/\delta} \frac{\delta t^2}{|z|^4 + t^2}\, dt \lesssim \int_0^{|z|^2} \frac{\delta t^2}{|z|^4}\, dt + \int_{|z|^2}^{1/\delta} \frac{\delta t^2}{t^2}\, dt \lesssim \delta \frac{|z|^6}{|z|^4} + 1 \lesssim 1,
$$

while for $|z|^2 > 1/\delta$, we have

$$
\int_0^{1/\delta} \frac{\delta t^2}{|z|^4 + t^2}\, dt \lesssim \int_0^{1/\delta} \frac{\delta t^2}{|z|^4}\, dt \lesssim \delta \frac{(1/\delta)^3}{|z|^4} \lesssim 1.
$$

Altogether we have $\left| \int_{\mathbb{R}} K(z,t)\phi_1(\delta t)\, dt \right| \lesssim |z|^{-2n}$ as required.

The second cancellation condition in (2) uses the assumption that Ω has mean zero on the sphere. For convenience we take $\beta = 0$. Then we have

$$
\left| \int_{\mathbb{C}^n} K(z,t)\phi_2(\delta z)\, dz \right| = \left| \int_{\mathbb{C}^n} \frac{\Omega(z)}{|z|^{2n}} \frac{1}{|z|^2 + it} \{\phi_2(\delta z) - \phi_2(0)\}\, dz \right|
$$

$$
\lesssim \delta \int_{\{|z| \leq 1/\delta\}} \frac{1}{|z|^{2n}} \frac{1}{|z|^2 + |t|} |z|\, dz
$$

$$
\lesssim \frac{\delta}{|t|} \int_0^{1/\delta} \frac{1}{r^{2n}} r(r^{2n-1}\, dr) \approx |t|^{-1},
$$

as required.

The third cancellation condition in (2) is handled similarly. We have

$$
\int_{\mathbb{H}^n} K(z,t)\phi_3(\delta_1 z, \delta_2 t)\, dz\, dt
$$

$$
= \int_{\mathbb{H}^n} \frac{\Omega(z)}{|z|^{2n}} \left\{ \frac{|z|^2}{|z|^4 + t^2} - \frac{it}{|z|^4 + t^2} \right\} \{\phi_3(\delta_1 z, \delta_2 t) - \phi_3(0, \delta_2 t)\}\, dz\, dt
$$

$$
= \int_{\mathbb{H}^n} \frac{\Omega(z)}{|z|^{2n}} \frac{|z|^2}{|z|^4 + t^2} \{\phi_3(\delta_1 z, \delta_2 t) - \phi_3(0, \delta_2 t)\}\, dz\, dt
$$

$$
- \int_{\mathbb{H}^n} \frac{\Omega(z)}{|z|^{2n}} \frac{it}{|z|^4 + t^2} \{\phi_3(\delta_1 z, \delta_2 t) - \phi_3(0, \delta_2 t) - \phi_3(\delta_1 z, 0) + \phi_3(0, 0)\}\, dz\, dt,
$$

and so

$$\left| \int_{\mathbb{H}^n} K(z,t)\phi_3(\delta_1 z, \delta_2 t)\, dz\, dt \right|$$

$$\lesssim \int_{|t| \le 1/\delta_2} \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n}} \frac{|z|^2}{|z|^4 + t^2} \delta_1 |z|\, dz\, dt + \int_{|t| \le 1/\delta_2} \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n}} \frac{|t|}{|z|^4 + t^2} \delta_1 |z| \delta_2 |t|\, dz\, dt = \mathrm{I} + \mathrm{II}.$$

Now if $1/\delta_2 \le |z|^2$, then

$$\mathrm{I} \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-3}} \left\{ \int_0^{1/\delta_2} \frac{1}{|z|^4}\, dt \right\} dz \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-1}}\, dz \approx \delta_1 \int_0^{1/\delta_1} dr = 1,$$

while if $1/\delta_2 > |z|^2$, then

$$\mathrm{I} \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-3}} \left\{ \int_0^{|z|^2} \frac{1}{|z|^4}\, dt + \int_{|z|^2}^{1/\delta_2} \frac{1}{t^2}\, dt \right\} dz \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-1}}\, dz \approx 1.$$

Finally, we have

$$\mathrm{II} \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-1}} \left\{ \delta_2 \int_{|t| \le 1/\delta_2} \frac{t^2}{|z|^4 + t^2}\, dt \right\} dz \lesssim \delta_1 \int_{|z| \le 1/\delta_1} \frac{1}{|z|^{2n-1}}\, dz \approx 1.$$

Part III. Appendix

Here in the appendix, we construct a *flag* dyadic decompositon of the Heisenberg group using the tiling theorem of Strichartz. See [Han et al. 2012] for an approach that generalizes to certain products of spaces of homogeneous type.

10. The Heisenberg grid

Let $\mathbb{H}^n = \mathbb{C}^n \times \mathbb{R}$ be the Heisenberg group with group multiplication

$$(\zeta, t) \cdot (\eta, s) = (\zeta + \eta, t + s + 2\operatorname{Im}(\zeta \cdot \bar{\eta})), \quad (\zeta, t), (\eta, s) \in \mathbb{C}^n \times \mathbb{R}.$$

Note that $(\eta, s)^{-1} = (-\eta, -s)$. Relative to this multiplication, we define the dilation

$$\delta_\lambda(\zeta, t) = (\lambda \zeta, \lambda^2 t),$$

and its corresponding "norm" on $\mathbb{H}^n$ by

$$\rho(\zeta, t) = \sqrt[4]{|\zeta|^4 + t^2}.$$

Then we define a symmetric quasimetric d on $\mathbb{H}^n$ by

$$d((\zeta, t), (\eta, s)) = \rho((\zeta, t) \cdot (\eta, s)^{-1}),$$

and note that

$$d(\delta_\lambda(\zeta, t), \delta_\lambda(\eta, s)) = \lambda d((\zeta, t), (\eta, s)).$$

The center of the group $\mathbb{H}^n$ is

$$\mathscr{L}^n = \{(\zeta, t) \in \mathbb{H}^n : \zeta = 0\},$$

which is isomorphic to the abelian group $\mathbb{R}$. The quotient group $\mathbb{Q}^n = \mathbb{H}^n / \mathscr{L}^n$ consists of equivalence classes $[(\zeta, t)]$ such that $[(\zeta, t)] = [(\eta, s)]$ if and only if

$$(\zeta, t) \cdot (\eta, s)^{-1} \in \mathscr{L}^n, \quad \text{that is, } \zeta = \eta.$$

Thus we may identify $\mathbb{Q}^n$ with $\mathbb{C}^n$ as abelian groups. Thus we see that $\mathbb{H}^n = \mathbb{C}^n \otimes_{\text{twist}} \mathbb{R}$ is a twisted group product of the abelian groups $\mathbb{C}^n$ and $\mathbb{R}$ with bihomomorphism $\beta(z, w) = 2\operatorname{Im}(z \cdot \bar{w})$. See the appendix for a discussion of this notion of twisted group product.

Now we apply the usual dyadic decomposition to the quotient metric space $\mathbb{Q}^n = \mathbb{C}^n$ to obtain a grid of "almost balls" (which are actually cubes here)

$$\{I\}_{I \text{ dyadic}} = \{I_\alpha^j\}_{j \in \mathbb{Z} \text{ and } \alpha \in 2^j \mathbb{Z}^{2n}},$$

where $I_0^j = [0, 2^j)^{2n}$ and $I_\alpha^j = I_0^j + \alpha$ for $j \in \mathbb{Z}$ and $\alpha \in 2^j \mathbb{Z}^{2n}$, so that $\ell(I_\alpha^j) = 2^j$. By a grid of almost balls we mean that the sets I_α^j decompose $\mathbb{C}^n$ at each scale 2^j, are almost balls, and are nested at differing scales; that is, there are positive constants C_1, C_2 and points $c_{I_\alpha^j} \in I_\alpha^j$ such that

$$\begin{aligned}
\mathbb{C}^n &= \dot{\bigcup} I_\alpha^j && j \in \mathbb{Z}, \\
B(c_{I_\alpha^j}, C_1 2^j) &\subset I_\alpha^j \subset B(c_{I_\alpha^j}, C_2 2^j) && j \in \mathbb{Z}, \alpha \in 2^j \mathbb{Z}^{2n}, \\
I_{\alpha'}^{j'} &\subset I_\alpha^j, \quad I_\alpha^j \subset I_{\alpha'}^{j'} \quad \text{or} \quad I_{\alpha'}^{j'} = I_\alpha^j.
\end{aligned} \tag{10-1}$$

Here we can take c_I to be the center of the cube I, and $C_1 = 1/2$, $C_2 = \sqrt{2n}/2 = \sqrt{n/2}$. We also have the usual dyadic grid $\{J_\tau^k\}_{k \in \mathbb{Z} \text{ and } \tau \in 2^k \mathbb{Z}}$ for $\mathbb{R}$, where $J_0^k = [0, 2^k)$ and $I_\tau^k = I_0^k + \tau$ for $k \in \mathbb{Z}$ and $\tau \in 2^k \mathbb{Z}$.

In order to use these grids to construct a "product-like" grid for $\mathbb{H}^n$, we must take into account the twisted structure of the product $\mathbb{H}^n = \mathbb{C}^n \otimes_{\text{twist}} \mathbb{R}$. Here is our theorem on the existence of a twisted grid for $\mathbb{H}^n$.

Theorem 68. *There is a positive integer m and positive constants C_1, C_2, such that, for each $j \in m\mathbb{Z}$ and*

$$(\alpha, \tau) \in K_j \equiv 2^j \mathbb{Z}^{2n} \times 2^{2j} \mathbb{Z},$$

there are subsets $\mathscr{S}_{j,\alpha,\tau}$ of $\mathbb{H}^n$ satisfying

$$\begin{aligned}
\mathbb{H}^n &= \dot{\bigcup_{(\alpha,\tau) \in K_j}} \mathscr{S}_{j,\alpha,\tau}, && \text{for each } j \in m\mathbb{Z}, \\
P_{\mathbb{C}^n} \mathscr{S}_{j,\alpha,\tau} &= I_\alpha^j, && j \in m\mathbb{Z}, (\alpha, \tau) \in K_j, \\
B_d(c_{j,\alpha,\tau}, C_1 2^j) &\subset \mathscr{S}_{j,\alpha,\tau} \subset B_d(c_{j,\alpha,\tau}, C_2 2^j), && j \in m\mathbb{Z}, (\alpha, \tau) \in K_j, \\
\mathscr{S}_{j,\alpha,\tau} &\subset \mathscr{S}_{j',\alpha',\tau'}, \quad \mathscr{S}_{j',\alpha',\tau'} \subset \mathscr{S}_{j,\alpha,\tau} \quad \text{or} \quad \mathscr{S}_{j,\alpha,\tau} \cap \mathscr{S}_{j',\alpha',\tau'} = \phi, \\
c_{j,\alpha,\tau} &= \left(P_{j,\alpha}, \tau + \tfrac{1}{2} 2^{2j}\right),
\end{aligned} \tag{10-2}$$

where $P_{j,\alpha} = c_{I_\alpha^j}$ and $P_{\mathbb{C}^n}$ denotes orthogonal projection of $\mathbb{H}^n$ onto $\mathbb{C}^n$.

Thus at each dyadic scale 2^j with $j \in m\mathbb{Z}$, we have a pairwise disjoint decomposition of $\mathbb{H}^n$ into sets $\mathscr{S}_{j,\alpha,\tau}$ that are almost Heisenberg balls of radius 2^j. These decompositions are nested, and moreover are *product-like* in the sense that the sets $\mathscr{S}_{j,\alpha,\tau}$ project onto the usual dyadic grid in the factor $\mathbb{C}^n$, and have centers $c_{j,\alpha,\tau} = \left(P_{j,\alpha}, \tau + \frac{1}{2}2^{2j}\right)$ that for each j form a product set indexed by $K_j \equiv 2^j\mathbb{Z}^{2n} \times 2^{2j}\mathbb{Z}$ and satisfy

$$|c_{j,\alpha,\tau} - c_{j,\alpha',\tau}| = 2^j \quad \text{and} \quad |c_{j,\alpha,\tau} - c_{j,\alpha,\tau'}| = 2^{2j},$$

if α and α' are neighbors in $2^j\mathbb{Z}^{2n}$ and if τ and τ' are neighbors in $2^{2j}\mathbb{Z}$.

Theorem 68 follows easily from the theory of self-similar tilings (neatly stacked over dyadic cubes) in [Strichartz 1992]. An excellent source for this material is [Tyson 2008, pp. 39–42]. See [Han et al. 2012] for more detail.

11. Rectangles in the Heisenberg group

Recall from Theorem 68 that at each dyadic scale 2^j with $j \in m\mathbb{Z}$ there is a pairwise disjoint decomposition of $\mathbb{H}^n$ into sets $\mathscr{S}_{j,\alpha,\tau}$ that are "almost Heisenberg ball" of radius 2^j. We will refer to these sets as dyadic cubes at scale 2^j. These decompositions are nested, and moreover are *product-like* in the sense that the cubes $\mathscr{S}_{j,\alpha,\tau}$ project onto I_α^j in the usual dyadic grid in the factor $\mathbb{C}^n$, and have centers $c_{j,\alpha,\tau} = \left(P_{j,\alpha}, \tau + \frac{1}{2}2^{2j}\right)$ that, for each j, form a product set indexed by $K_j \equiv 2^j\mathbb{Z}^{2n} \times 2^{2j}\mathbb{Z}$ and satisfy

$$|c_{j,\alpha,\tau} - c_{j,\alpha',\tau}| = 2^j \quad \text{and} \quad |c_{j,\alpha,\tau} - c_{j,\alpha,\tau'}| = 2^{2j},$$

if α and α' are neighbors in $2^j\mathbb{Z}^{2n}$ and if τ and τ' are neighbors in $2^{2j}\mathbb{Z}$.

We now define vertical and horizontal dyadic rectangles relative to this decomposition into dyadic cubes. The analogy with dyadic rectangles in the plane $\mathbb{R}^2$ that we are pursuing here is that a dyadic rectangle $I = I_1 \times I_2$ in the plane is vertical if $|I_2| \geq |I_1|$, and is horizontal if $|I_1| \geq |I_2|$ (and both if and only if I is a dyadic square). If we consider the grid of dyadic cubes $\{\mathscr{S}_{j,\alpha,\tau}\}$ in $\mathbb{H}^n$ in place of the grid of dyadic squares in $\mathbb{R}^2$, we are led to the following definition.

Definition 69. Let $j, k \in m\mathbb{Z}$, with $j \leq k$, and let $\mathscr{S}_{j,\alpha,\tau}$ and $\mathscr{S}_{k,\beta,\upsilon}$ be dyadic cubes in $\mathbb{H}^n$ with $\mathscr{S}_{j,\alpha,\tau} \subset \mathscr{S}_{k,\beta,\upsilon}$. The set

$$\mathscr{R}(\text{ver}) = \mathscr{R}^{\mathscr{S}_{k,\beta,\upsilon}}_{\mathscr{S}_{j,\alpha,\tau}}(\text{ver}) = \bigcup \{\mathscr{S}_{j,\alpha,\tau'} : \mathscr{S}_{j,\alpha,\tau'} \subset \mathscr{S}_{k,\beta,\upsilon}\}$$

will be referred to as a *vertical dyadic rectangle*, or, more precisely, the vertical dyadic rectangle in $\mathscr{S}_{k,\beta,\upsilon}$ containing $\mathscr{S}_{j,\alpha,\tau}$. We define the *base* of the rectangle $\mathscr{R}(\text{ver})$ to be the dyadic cube I_α^j in $\mathbb{C}^n$, and we define the *cobase* of the rectangle $\mathscr{R}(\text{ver})$ to be the dyadic interval J_υ^{2k} in $\mathbb{R}$. We say the rectangle $\mathscr{R}(\text{ver})$ has *width* 2^j and *height* 2^{2k}. Similarly, the set

$$\mathscr{R}(\text{hor}) = \mathscr{R}^{\mathscr{S}_{k,\beta,\upsilon}}_{\mathscr{S}_{j,\alpha,\tau}}(\text{hor}) = \bigcup \{\mathscr{S}_{j,\alpha',\tau} : \mathscr{S}_{j,\alpha',\tau} \subset \mathscr{S}_{k,\beta,\upsilon}\}$$

will be referred to as a *horizontal dyadic rectangle*, or, more precisely, the horizontal dyadic rectangle in $\mathscr{S}_{k,\beta,\upsilon}$ containing $\mathscr{S}_{j,\alpha,\tau}$. We define the *base* of the rectangle $\mathscr{R}(\text{hor})$ to be the dyadic cube I_β^k in $\mathbb{C}^n$, and

we define the *cobase* of the rectangle $\mathcal{R}(\text{ver})$ to be the dyadic interval J_τ^{2j} in $\mathbb{R}$. We say the rectangle $\mathcal{R}(\text{hor})$ has *width* 2^k and *height* 2^{2j}.

We will usually write just $\mathcal{R}$ to denote a dyadic rectangle that is either vertical or horizontal. Note that a dyadic rectangle $\mathcal{R}$ is both vertical and horizontal if and only if $\mathcal{R}$ is a dyadic cube $\mathcal{S}_{j,\alpha,\tau}$. Finally note that $\mathcal{R}_{\mathcal{S}_{j,\alpha,\tau}}^{\mathcal{S}_{k,\beta,\upsilon}}(\text{ver})$ can be thought of as a Heisenberg substitute for the Euclidean rectangle $I_\alpha^j \times J_\upsilon^{2k}$ in $\mathbb{H}^n$ with width 2^j and height 2^{2k}, and that $\mathcal{R}_{\mathcal{S}_{j,\alpha,\tau}}^{\mathcal{S}_{k,\beta,\upsilon}}(\text{hor})$ can be thought of as a Heisenberg substitute for the Euclidean rectangle $I_\beta^k \times J_\tau^{2j}$ in $\mathbb{H}^n$ with width 2^k and height 2^{2j}. The vertical Heisenberg rectangles are constructed by stacking Heisenberg cubes neatly on top of each other, while the horizontal Heisenberg rectangles are constructed by placing Heisenberg cubes next to each other, although the placement is far from neat.

Remark 70. In applications to operators with flag kernels, or more generally a semiproduct structure, it is appropriate to restrict attention to the set of *vertical* dyadic rectangles.

References

[Carleson 1974] L. Carleson, "A counterexample for measures bounded on H^p for the bidisc", technical report, Institut Mittag-Leffler, 1974.

[Chang 1979] S.-Y. A. Chang, "Carleson measure on the bi-disc", *Ann. of Math.* (2) **109**:3 (1979), 613–620. MR 80j:32009 Zbl 0401.28004

[Chang and Fefferman 1980] S.-Y. A. Chang and R. Fefferman, "A continuous version of duality of H^1 with BMO on the bidisc", *Ann. of Math.* (2) **112**:1 (1980), 179–201. MR 82a:32009 Zbl 0451.42014

[Chang and Fefferman 1982] S.-Y. A. Chang and R. Fefferman, "The Calderón–Zygmund decomposition on product domains", *Amer. J. Math.* **104**:3 (1982), 455–468. MR 84a:42028 Zbl 0513.42019

[Chang and Fefferman 1985] S.-Y. A. Chang and R. Fefferman, "Some recent developments in Fourier analysis and H^p-theory on product domains", *Bull. Amer. Math. Soc.* (N.S.) **12**:1 (1985), 1–43. MR 86g:42038 Zbl 0557.42007

[Coifman and Weiss 1976] R. R. Coifman and G. Weiss, *Transference methods in analysis*, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics **31**, American Mathematical Society, Providence, 1976. MR 58 #2019 Zbl 0377.43001

[Fefferman 1986] R. Fefferman, "Multiparameter Fourier analysis", pp. 47–130 in *Beijing lectures in harmonic analysis* (Beijing, 1984), edited by E. M. Stein, Ann. of Math. Stud. **112**, Princeton Univ. Press, Princeton, 1986. MR 89a:42001 Zbl 0632.42014

[Fefferman 1987] R. Fefferman, "Harmonic analysis on product spaces", *Ann. of Math.* (2) **126**:1 (1987), 109–130. MR 90e:42030 Zbl 0644.42017

[Fefferman 1999] R. A. Fefferman, "Multiparameter Calderón–Zygmund theory", pp. 207–221 in *Harmonic analysis and partial differential equations* (Chicago, 1996), edited by M. Christ et al., Univ. Chicago Press, Chicago, 1999. MR 2001c:42011 Zbl 0952.42009

[Fefferman and Stein 1972] C. Fefferman and E. M. Stein, "H^p spaces of several variables", *Acta Math.* **129**:3-4 (1972), 137–193. MR 56 #6263 Zbl 0257.46078

[Fefferman and Stein 1982] R. Fefferman and E. M. Stein, "Singular integrals on product spaces", *Adv. in Math.* **45**:2 (1982), 117–143. MR 84d:42023 Zbl 0517.42024

[Ferguson and Lacey 2002] S. H. Ferguson and M. T. Lacey, "A characterization of product BMO by commutators", *Acta Math.* **189**:2 (2002), 143–160. MR 2004e:42026 Zbl 1039.47022

[Frazier and Jawerth 1990] M. Frazier and B. Jawerth, "A discrete transform and decompositions of distribution spaces", *J. Funct. Anal.* **93**:1 (1990), 34–170. MR 92a:46042 Zbl 0716.46031

[Geller and Mayeli 2006] D. Geller and A. Mayeli, "Continuous wavelets and frames on stratified Lie groups, I", *J. Fourier Anal. Appl.* **12**:5 (2006), 543–579. MR 2007g:42055 Zbl 1105.42029

[Gilbert et al. 2002] J. Gilbert, Y. Han, J. Hogan, J. Lakey, D. Weiland, and G. Weiss, "Smooth molecular decompositions of functions and singular integral operators", *Mem. Am. Math. Soc.* **742** (2002), 74.

[Gundy and Stein 1979] R. F. Gundy and E. M. Stein, "H^p theory for the poly-disc", *Proc. Nat. Acad. Sci. U.S.A.* **76**:3 (1979), 1026–1029. MR 80j:32012 Zbl 0405.32002

[Han 1994] Y. S. Han, "Calderón-type reproducing formula and the Tb theorem", *Rev. Mat. Iberoamericana* **10**:1 (1994), 51–91. MR 95h:42020 Zbl 0797.42009

[Han 1998] Y.-S. Han, "Plancherel–Pólya type inequality on spaces of homogeneous type and its applications", *Proc. Amer. Math. Soc.* **126**:11 (1998), 3315–3327. MR 99a:42010 Zbl 0920.42011

[Han and Lu 2008] Y. Han and G. Lu, "Discrete Littlewood–Paley–Stein theory and multi-parameter Hardy spaces associated with flag singular integrals", preprint, 2008. arXiv 0801.1701

[Han and Lu 2010] Y. Han and G. Lu, "Some recent works on multiparameter Hardy space theory and discrete Littlewood–Paley analysis", pp. 99–191 in *Trends in partial differential equations*, Adv. Lect. Math. (ALM) **10**, International Press, 2010. MR 2011e:42043 Zbl 1201.42016

[Han et al. 2012] Y. Han, G. Lu, and E. Sawyer, "Flag Hardy spaces and Marcinkiewicz multipliers on the Heisenberg group: an expanded version", preprint, 2012. arXiv 1208.2484

[Han et al. 2013a] Y. Han, J. Li, and G. Lu, "Multiparameter Hardy space theory on Carnot–Carathéodory spaces and product spaces of homogeneous type", *Trans. Amer. Math. Soc.* **365**:1 (2013), 319–360. MR 2984061 Zbl 1275.42035

[Han et al. 2013b] Y. Han, C. Lin, G. Lu, Z. Ruan, and E. T. Sawyer, "Hardy spaces associated with different homogeneities and boundedness of composition operators", *Rev. Mat. Iberoam.* **29**:4 (2013), 1127–1157. MR 3148598 Zbl 1291.42018

[Jessen et al. 1935] B. Jessen, J. Marcinkiewicz, and A. Zygmund, "Note on the differentiability of multiple integrals", *Fundam. Math.* **25** (1935), 217–234. Zbl 0012.05901

[Journé 1985] J.-L. Journé, "Calderón–Zygmund operators on product spaces", *Rev. Mat. Iberoamericana* **1**:3 (1985), 55–91. MR 88d:42028 Zbl 0634.42015

[Journé 1986] J.-L. Journé, "A covering lemma for product spaces", *Proc. Amer. Math. Soc.* **96**:4 (1986), 593–598. MR 87g:42028 Zbl 0594.42015

[Journé 1988] J.-L. Journé, "Two problems of Calderón–Zygmund theory on product-spaces", *Ann. Inst. Fourier* (*Grenoble*) **38**:1 (1988), 111–132. MR 90b:42031 Zbl 0638.47026

[Lacey and Terwilliger 2005] M. Lacey and E. Terwilliger, "Little Hankel operators and product BMO", lecture presented at harmonic analysis and its applications, 2005, Available at http://people.math.gatech.edu/~lacey/talks/hankel.pdf.

[Müller et al. 1995] D. Müller, F. Ricci, and E. M. Stein, "Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups, I", *Invent. Math.* **119**:2 (1995), 199–233. MR 96b:43005 Zbl 0857.43012

[Müller et al. 1996] D. Müller, F. Ricci, and E. M. Stein, "Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups, II", *Math. Z.* **221**:2 (1996), 267–291. MR 97c:43007 Zbl 0863.43001

[Nagel and Stein 2004] A. Nagel and E. M. Stein, "On the product theory of singular integrals", *Rev. Mat. Iberoamericana* **20**:2 (2004), 531–561. MR 2006i:42023 Zbl 1057.42016

[Nagel et al. 2001] A. Nagel, F. Ricci, and E. M. Stein, "Singular integrals with flag kernels and analysis on quadratic CR manifolds", *J. Funct. Anal.* **181**:1 (2001), 29–118. MR 2001m:22018 Zbl 0974.22007

[Nagel et al. 2012] A. Nagel, F. Ricci, E. Stein, and S. Wainger, "Singular integrals with flag kernels on homogeneous groups, I", *Rev. Mat. Iberoam.* **28**:3 (2012), 631–722. MR 2949616 Zbl 1253.42009

[Pipher 1986] J. Pipher, "Journé's covering lemma and its extension to higher dimensions", *Duke Math. J.* **53**:3 (1986), 683–690. MR 88a:42019 Zbl 0645.42018

[Plancherel and Pólya 1937] M. Plancherel and G. Pólya, "Fonctieres entières et intégrales de Fourier multiples", *Comment. Math. Helv.* **9** (1937), 224–248. Zbl 0016.36004

[Pólya 1936] G. Pólya, "Fonctions entières et intégrales de Fourier multiples", *Comment. Math. Helv.* **9**:1 (1936), 224–248. MR 1509557 Zbl 0016.36004

[Strichartz 1992] R. S. Strichartz, "Self-similarity on nilpotent Lie groups", pp. 123–157 in *Geometric analysis* (Philadelphia, 1991), edited by E. L. Grinberg, Contemp. Math. **140**, Amer. Math. Soc., Providence, 1992. MR 94e:43011 Zbl 0797.43004

[Torres 1991] R. H. Torres, "Boundedness results for operators with singular kernels on distribution spaces", *Mem. Amer. Math. Soc.* **90**:442 (1991), viii+172. MR 91g:47044 Zbl 0737.47041

[Tyson 2008] J. T. Tyson, "Global conformal Assouad dimension in the Heisenberg group", *Conform. Geom. Dyn.* **12** (2008), 32–57. MR 2009a:30052 Zbl 1133.30323

YONGSHENG HAN: hanyong@mail.auburn.edu
Department of Mathematics, Auburn University, Auburn, AL 36849, United States

GUOZHEN LU: gzlu@math.wayne.edu
Department of Mathematics, Wayne State University, 656 W. Kirby Street, Detroit, MI 48202, United States

ERIC SAWYER: sawyer@mcmaster.ca
Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada

■msp

RIGIDITY OF EQUALITY CASES IN STEINER'S PERIMETER INEQUALITY

Filippo Cagnetti, Maria Colombo, Guido De Philippis and Francesco Maggi

Dedicated to Nicola Fusco, for his mentorship

Characterization results for equality cases and for rigidity of equality cases in Steiner's perimeter inequality are presented. (By rigidity, we mean the situation when all equality cases are vertical translations of the Steiner symmetral under consideration.) We achieve this through the introduction of a suitable measure-theoretic notion of connectedness and a fine analysis of barycenter functions for sets of finite perimeter having segments as orthogonal sections with respect to a hyperplane.

1. Introduction

1A. *Overview.* Steiner symmetrization is a classical and powerful tool in the analysis of geometric variational problems. Indeed, while volume is preserved under Steiner symmetrization, other relevant geometric quantities, like diameter or perimeter, behave monotonically. In particular, Steiner's perimeter inequality asserts the crucial fact that perimeter is decreased by Steiner symmetrization, a property that, in turn, lies at the heart of a well-known proof of the Euclidean isoperimetric theorem; see [De Giorgi 1958]. In the seminal paper [Chlebík et al. 2005], which we briefly review in Section 1B, Chlebík, Cianchi and Fusco discuss Steiner's inequality in the natural framework of sets of finite perimeter, and provide a sufficient condition for the rigidity of equality cases. By *rigidity of equality cases* we mean that situation when the only sets achieving equality in Steiner's inequality are obtained as translations of the Steiner symmetral. Roughly speaking, the sufficient condition for rigidity found in [Chlebík et al. 2005] amounts to requiring that the Steiner symmetral has "no vertical boundary" and "no vanishing sections". While simple examples show that rigidity may indeed fail if one of these two assumptions is dropped, it is likewise easy to construct polyhedral Steiner symmetrals such that rigidity holds and *both* these conditions

are violated. In particular, the problem of a geometric characterization of rigidity of equality cases in Steiner's inequality was left open in [Chlebík et al. 2005], even in the fundamental case of polyhedra.

In the recent paper [Cagnetti et al. 2013], we have fully addressed the rigidity problem in the case of Ehrhard's inequality for a Gaussian perimeter. Indeed, we obtain a *characterization* of rigidity, rather than a mere sufficient condition for it. A crucial step in proving (and, actually, formulating) this sharp result consists in the introduction of a measure-theoretic notion of connectedness, and, more precisely, of what it means for a Borel set to "disconnect" another Borel set; see Section 1C for more details.

In this paper, we aim to exploit these ideas in the study of Steiner's perimeter inequality. In order to achieve this goal we shall need a sharp description of the properties of the barycenter function of a set of finite perimeter having segments as orthogonal sections with respect to a hyperplane (Theorem 1.7). With these tools at hand, we completely characterize equality cases in Steiner's inequality in terms of properties of their barycenter functions (Theorem 1.9). Starting from this result, we obtain a general sufficient condition for rigidity (Theorem 1.11), and we show that, if the slice length function is of special bounded variation with locally finite jump set, then equality cases are necessarily obtained by at most countably many vertical translations of "chunks" of the Steiner symmetral (Theorem 1.13); see Section 1D.

In Section 1E, we introduce several characterizations of rigidity. In Theorem 1.16 we provide *two* geometric characterizations of rigidity under the "no vertical boundary" assumption considered in [Chlebík et al. 2005]. In Theorem 1.20 we characterize rigidity in the case when the Steiner symmetral is a generalized polyhedron. (Here, the generalization of the usual notion of polyhedron consists in replacing affine functions over bounded polygons with $W^{1,1}$-functions over sets of finite perimeter and volume.) We then characterize rigidity when the slice length function is of special bounded variation with locally finite jump set, by introducing a condition we call the *mismatched stairway property* (Theorem 1.29). Finally, in Theorem 1.30, we prove two characterizations of rigidity in the planar setting.

By building on the results and methods introduced in this paper, it is of course possible to analyze the rigidity problem for Steiner perimeter inequalities in higher codimensions. Although it would have been natural to discuss these issues here, the already considerable length and technical complexity of the present paper suggested we do this in a separate forthcoming paper.

1B. *The Steiner inequality and the rigidity problem.* We begin by recalling the definition of Steiner symmetrization and the main result from [Chlebík et al. 2005]. In doing so, we shall refer to some concepts from the theory of sets of finite perimeter and functions of bounded variation (that are summarized in Section 2B), and we shall fix a minimal set of notation used through the rest of the paper. We decompose $\mathbb{R}^n$, $n \geq 2$, as the Cartesian product $\mathbb{R}^{n-1} \times \mathbb{R}$, denoting by $\boldsymbol{p} : \mathbb{R}^n \to \mathbb{R}^{n-1}$ and $\boldsymbol{q} : \mathbb{R}^n \to \mathbb{R}$ the horizontal and vertical projections, so that $x = (\boldsymbol{p}x, \boldsymbol{q}x)$ with $\boldsymbol{p}x = (x_1, \ldots, x_{n-1})$, $\boldsymbol{q}x = x_n$ for every $x \in \mathbb{R}^n$. Given $E \subset \mathbb{R}^n$ we denote by E_z the vertical section of E with respect to $z \in \mathbb{R}^{n-1}$, that is, we set

$$E_z = \{t \in \mathbb{R} : (z, t) \in E\}.$$

Moreover, given a function $v : \mathbb{R}^{n-1} \to [0, \infty)$, we say that E is v-*distributed* if

$$v(z) = \mathcal{H}^1(E_z) \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } z \in \mathbb{R}^{n-1}.$$

(Here, $\mathcal{H}^k(S)$ stands for the k-dimensional Hausdorff measure on the Euclidean space containing the set S under consideration.) Among all v-distributed sets, we denote by $F[v]$ the (only) one that is symmetric by reflection with respect to $\{\boldsymbol{q}x = 0\}$, and whose vertical sections are segments, that is, we set

$$F[v] = \{x \in \mathbb{R}^n : |\boldsymbol{q}x| < \tfrac{1}{2}v(\boldsymbol{p}x)\}.$$

If E is a v-distributed set, then the set $F[v]$ is the *Steiner symmetral* of E, and is usually denoted as E^s. (Our notation reflects the fact that, in addressing the structure of equality cases, we are more concerned with properties of v rather than with the properties of a particular v-distributed set.) The set $F[v]$ has finite volume if and only if $v \in L^1(\mathbb{R}^{n-1})$, and it is of finite perimeter if and only if $v \in BV(\mathbb{R}^{n-1})$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$; see Proposition 3.2. Denoting by $P(E; A)$ the relative perimeter of E with respect to the Borel set $A \subset \mathbb{R}^n$ (so that, for example, $P(E; A) = \mathcal{H}^{n-1}(A \cap \partial E)$ if E is an open set with Lipschitz boundary in $\mathbb{R}^n$), the *Steiner perimeter inequality* implies that, if E is a v-distributed set of finite perimeter, then

$$P(E; G \times \mathbb{R}) \geq P(F[v]; G \times \mathbb{R}) \quad \text{for every Borel set } G \subset \mathbb{R}^{n-1}. \tag{1-1}$$

Inequality (1-1) was first proved in this generality by De Giorgi [1958], in the course of his proof of the Euclidean isoperimetric theorem for sets of finite perimeter. Indeed, an important step in his argument consists in showing that if a set E satisfies (1-1) with equality, then, for $\mathcal{H}^{n-1}$-a.e. $z \in G$, the vertical section E_z is $\mathcal{H}^1$-equivalent to a segment; see [Maggi 2012, Chapter 14]. The study of equality cases in Steiner's inequality was then resumed by Chlebík et al. [2005]. We now recall two important results from their paper. The first theorem, which is easily deduced by means of [Chlebík et al. 2005, Theorem 1.1, Proposition 4.2], completes De Giorgi's analysis of necessary conditions for equality, and, in turn, provides a characterization of equality cases whenever $\partial^* E$ has no vertical parts. Given a Borel set $G \subset \mathbb{R}^{n-1}$, we set

$$\mathcal{M}_G(v) = \{E \subset \mathbb{R}^n : E \text{ v-distributed and } P(E; G \times \mathbb{R}) = P(F[v]; G \times \mathbb{R})\} \tag{1-2}$$

to denote the family of sets achieving equality in (1-1), and simply set $\mathcal{M}(v) = \mathcal{M}_{\mathbb{R}^{n-1}}(v)$.

Theorem A [Chlebík et al. 2005]. *Let $v \in BV(\mathbb{R}^{n-1})$ and let E be a v-distributed set of finite perimeter. If $E \in \mathcal{M}_G(v)$ then, for $\mathcal{H}^{n-1}$-a.e. $z \in G$, E_z is $\mathcal{H}^1$-equivalent to a segment (t^-, t^+), with $(z, t^+), (z, t^-) \in \partial^* E$, $\boldsymbol{p}v_E(z, t^+) = \boldsymbol{p}v_E(z, t^-)$, and $\boldsymbol{q}v_E(z, t^+) = -\boldsymbol{q}v_E(z, t^-)$.*

The converse implication holds provided $\partial^* E$ has no vertical parts above G, that is,

$$\mathcal{H}^{n-1}\big(\{x \in \partial^* E \cap (G \times \mathbb{R}) : \boldsymbol{q}v_E(x) = 0\}\big) = 0, \tag{1-3}$$

where $\partial^* E$ denotes the reduced boundary of E, while v_E is the measure-theoretic outer unit normal of E; see Section 2B.

The second main result, from [Chlebík et al. 2005, Theorem 1.3], provides a sufficient condition for the rigidity of equality cases in Steiner's inequality over an open connected set. Note indeed that some assumptions are needed in order to expect rigidity; see Figure 1.
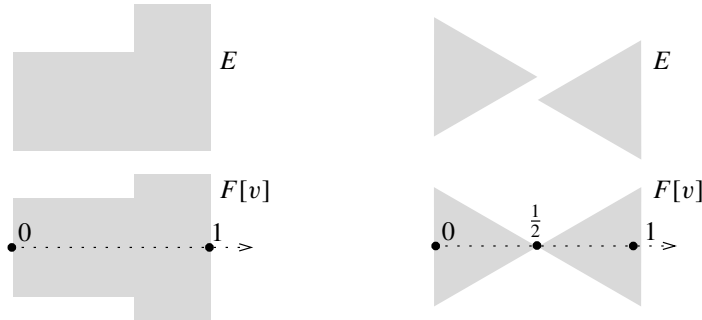
Figure 1. Left: $\partial^* F[v]$ has vertical parts over $\Omega = (0, 1)$ and (1-6) does not hold. Right: $\partial^* F[v]$ has no vertical parts over $\Omega = (0, 1)$, but (1-5) fails (indeed, $0 = v^\vee(\frac{1}{2}) = v^\wedge(\frac{1}{2})$).

Theorem B [Chlebík et al. 2005]. *If $v \in BV(\mathbb{R}^{n-1})$, $\Omega \subset \mathbb{R}^{n-1}$ is an open connected set with $\mathcal{H}^{n-1}(\Omega) < \infty$, and*

$$D^s v \llcorner \Omega = 0, \tag{1-4}$$

$$v^\wedge > 0 \quad \mathcal{H}^{n-2}\text{-a.e. on } \Omega, \tag{1-5}$$

then for every $E \in \mathcal{M}_\Omega(v)$ we have

$$\mathcal{H}^n\big(\big(E \Delta(t e_n + F[v])\big) \cap (\Omega \times \mathbb{R})\big) = 0 \quad \text{for some } t \in \mathbb{R}. \tag{1-6}$$

Remark 1.1. Here, $D^s v$ stands for the singular part of the distributional derivative Dv of v, while $v^\wedge$ and $v^\vee$ denote the approximate lower and upper limits of v (so that if $v_1 = v_2$ a.e. on $\mathbb{R}^{n-1}$, then $v_1^\vee = v_2^\vee$ and $v_1^\wedge = v_2^\wedge$ *everywhere* on $\mathbb{R}^{n-1}$). We call $[v] = v^\vee - v^\wedge$ the jump of v, and define the approximate discontinuity set of v as $S_v = \{v^\vee > v^\wedge\} = \{[v] > 0\}$, so that S_v is countably $\mathcal{H}^{n-2}$-rectifiable, and there exists a Borel vector field $\nu_v : S_v \to S^{n-1}$ such that $D^s v = \nu_v[v]\mathcal{H}^{n-2}\llcorner S_v + D^c v$, where $D^c v$ stands for the Cantorian part of Dv. These concepts are reviewed in Sections 2A and 2B.

Remark 1.2. Assumption (1-4) is clearly equivalent to asking that $v \in W^{1,1}(\Omega)$ (so that $v^\wedge = v^\vee$ $\mathcal{H}^{n-2}$-a.e. on Ω), and, in turn, it is also equivalent to asking that $\partial^* F[v]$ have no vertical parts above Ω, that is — compare with (1-3) —

$$\mathcal{H}^{n-1}\big(\{x \in \partial^* F[v] \cap (\Omega \times \mathbb{R}) : \boldsymbol{q}\nu_{F[v]}(x) = 0\}\big) = 0; \tag{1-7}$$

see [Chlebík et al. 2005, Proposition 1.2] for a proof.

Remark 1.3. Although assuming the "no vertical parts" (1-4) and "no vanishing sections" (1-5) conditions appears natural in light of the examples sketched in Figure 1, it should be noted that these assumptions are far from being necessary for rigidity. For example, Figure 2 shows the case of a polyhedron in $\mathbb{R}^3$ such that (1-6) holds, but the "no vertical parts" condition fails. Similarly, in Figure 3, we have a polyhedron in $\mathbb{R}^3$ such that (1-6) and (1-4) hold, but such that (1-5) fails.
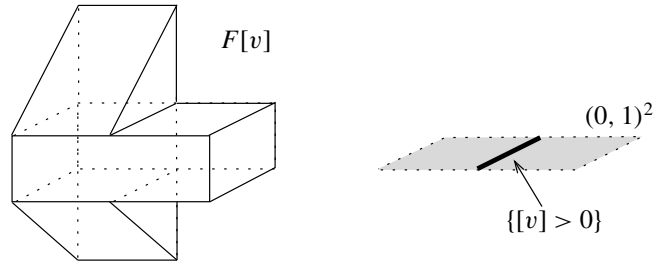
Figure 2. A polyhedron in $\mathbb{R}^3$ such that the rigidity condition (1-6) is satisfied (with $\Omega = (0, 1)^2$) but the "no vertical parts" condition fails.
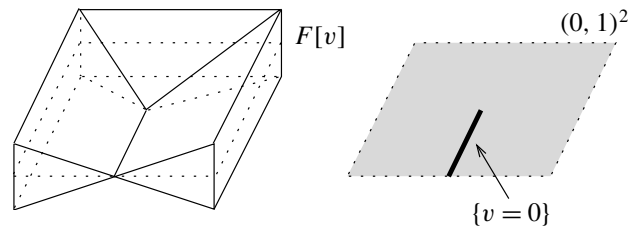
Figure 3. A polyhedron in $\mathbb{R}^3$ such that the rigidity condition (1-6) and the "no vertical parts" condition hold (with $\Omega = (0, 1)^2$), but the "no vanishing sections" condition fails.

1C. *Essential connectedness.* The examples discussed in Figure 1 and Remark 1.3 suggest that in order to characterize rigidity of equality cases in Steiner's inequality one should first make precise the sense in which the $(n-2)$-dimensional set $S_v = \{v^\wedge < v^\vee\}$ (contained in the projection of vertical boundaries) may disconnect the $(n-1)$-dimensional set $\{v > 0\}$ (that is, the projection of $F[v]$). In the study of rigidity of equality cases for Ehrhard's perimeter inequality — see [Cagnetti et al. 2013] — we have addressed this kind of question by introducing the following definition.

Definition 1.4. Let K and G be Borel sets in $\mathbb{R}^m$. One says that K *essentially disconnects* G if there exists a nontrivial Borel partition $\{G_+, G_-\}$ of G modulo $\mathcal{H}^m$ such that

$$\mathcal{H}^{m-1}((G^{(1)} \cap \partial^e G_+ \cap \partial^e G_-) \setminus K) = 0; \tag{1-8}$$

conversely, one says that K *does not essentially disconnect* G if, for every nontrivial Borel partition $\{G_+, G_-\}$ of G modulo $\mathcal{H}^m$,

$$\mathcal{H}^{m-1}((G^{(1)} \cap \partial^e G_+ \cap \partial^e G_-) \setminus K) > 0. \tag{1-9}$$

Finally, G is *essentially connected* if $\varnothing$ does not essentially disconnect G.

In the above definition, by a nontrivial Borel partition $\{G_+, G_-\}$ of G modulo $\mathcal{H}^m$ we mean that

$$\mathcal{H}^m(G_+ \cap G_-) = 0, \quad \mathcal{H}^m(G \Delta (G_+ \cup G_-)) = 0, \quad \mathcal{H}^m(G_+)\mathcal{H}^m(G_-) > 0.$$

Moreover, $\partial^e G$ denotes the essential boundary of G, which is defined as

$$\partial^e G = \mathbb{R}^m \setminus (G^{(0)} \cup G^{(1)}),$$
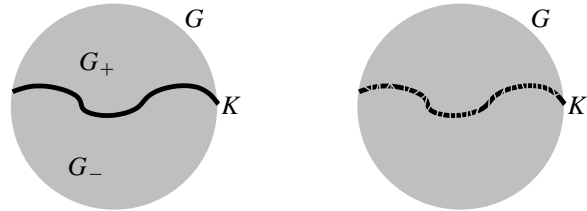
Figure 4. Left: G is a disk and K is a smooth curve that divides G in two open regions G_+ and G_-, in such a way that (1-8) holds: thus, K essentially disconnects G. Right: Let K' be obtained by removing some points from K. If we remove a set of length zero, that is, if $\mathcal{H}^1(K \setminus K') = 0$, then K' still essentially disconnects G (although $G \setminus K'$ may easily be topologically connected); if, instead, $\mathcal{H}^1(K \setminus K') > 0$, then K' does not essentially disconnect G, since (1-9) holds (with K' in place of K).

where $G^{(0)}$ and $G^{(1)}$ denote the sets of points of density 0 and 1 of G; see Section 2A.

Remark 1.5. If $\mathcal{H}^m(G \triangle G') = 0$ and $\mathcal{H}^{m-1}(K \triangle K') = 0$, then K essentially disconnects G if and only if K' essentially disconnects G'; see Figure 4.

Remark 1.6. We refer to [Cagnetti et al. 2013, Section 1.5] for more comments on the relation between this definition and the notions of indecomposable currents [Federer 1969, 4.2.25] and indecomposable sets of finite perimeter [Dolzmann and Müller 1995, Definition 2.11] or [Ambrosio et al. 2001, Section 4] used in geometric measure theory. We just recall here that a set of finite perimeter E is said to be *indecomposable* if $P(E) < P(E_+) + P(E_-)$ whenever $\{E_+, E_-\}$ is a nontrivial partition modulo $\mathcal{H}^n$ of E by sets of finite perimeter. Moreover, the latter inequality is equivalent to $\mathcal{H}^{n-1}(E^{(1)} \cap \partial^e E_+ \cap \partial^e E_-) > 0$. Let us also note that this measure-theoretic notion of connectedness is compatible with essential connectedness: indeed, as proved in [Cagnetti et al. 2013, Remark 2.3], a set of finite perimeter is indecomposable if and only if it is essentially connected. Nevertheless, when possible, we shall use the term indecomposable in place of the term essentially connected, in order to make immediate the identification of those statements and conditions whose formulation genuinely requires Definition 1.4.

1D. *Equality cases and barycenter functions.* With the notion of essential connectedness at hand we can easily conjecture several possible improvements of Theorem B. As it turns out, a fine analysis of the barycenter function for sets of finite perimeter with segments as sections is crucial in order to actually prove these results. Given a v-distributed set E, we define the barycenter function of E, $b_E : \mathbb{R}^{n-1} \to \mathbb{R}$, by setting, for every $z \in \mathbb{R}^{n-1}$,

$$b_E(z) = \frac{1}{v(z)} \int_{E_z} t \, d\mathcal{H}^1(t) \quad \text{if } v(z) > 0 \text{ and } \int_{E_z} t \, d\mathcal{H}^1(t) \in \mathbb{R}, \tag{1-10}$$

and $b_E(z) = 0$ otherwise. In general, b_E may only be a Lebesgue measurable function. When E has segments as sections and finite perimeter, the following theorem provides a degree of regularity for b_E that turns out to be sharp; see Remark 3.5. Note that the set where v vanishes is critical for the regularity of the barycenter, as implicitly expressed by (1-11).

Theorem 1.7. *If $v \in BV(\mathbb{R}^{n-1})$ and E is a v-distributed set of finite perimeter such that E_z is $\mathcal{H}^1$-equivalent to a segment for $\mathcal{H}^{n-1}$-a.e. $z \in \mathbb{R}^{n-1}$, then*

$$b_\delta = 1_{\{v > \delta\}} b_E \in GBV(\mathbb{R}^{n-1}) \tag{1-11}$$

for every $\delta > 0$ such that $\{v > \delta\}$ is a set of finite perimeter. Moreover, b_E is approximately differentiable $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, and for every Borel set $G \subset \{v^\wedge > 0\}$ we have the coarea formula

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \partial^e \{b_E > t\}) \, dt = \int_G |\nabla b_E| \, d\mathcal{H}^{n-1} + \int_{G \cap S_{b_E}} [b_E] \, d\mathcal{H}^{n-2} + |D^c b_E|^+(G), \tag{1-12}$$

where $|D^c b_E|^+$ is the Borel measure on $\mathbb{R}^{n-1}$ defined by

$$|D^c b_E|^+(G) = \lim_{\delta \to 0^+} |D^c b_\delta|(G) = \sup_{\delta > 0} |D^c b_\delta|(G) \quad \text{for all } G \subset \mathbb{R}^{n-1}. \tag{1-13}$$

Remark 1.8. Let us recall that $u \in GBV(\mathbb{R}^{n-1})$ if and only if $\tau_M(u) \in BV_{\text{loc}}(\mathbb{R}^{n-1})$ for every $M > 0$ (where $\tau_M(s) = \max\{-M, \min\{M, s\}\}$ for $s \in \mathbb{R}$), and that for every $u \in GBV(\mathbb{R}^{n-1})$ we can define a Borel measure $|D^c u|$ on $\mathbb{R}^{n-1}$ by setting

$$|D^c u|(G) = \lim_{M \to \infty} |D^c(\tau_M u)|(G) = \sup_{M > 0} |D^c(\tau_M u)|(G) \tag{1-14}$$

for every Borel set $G \subset \mathbb{R}^{n-1}$. (If $u \in BV(\mathbb{R}^{n-1})$, then the total variation of the Cantorian part of Du agrees with the measure defined in (1-14) on every Borel set.) The measures $|D^c b_\delta|$ appearing in (1-13) are thus defined by means of (1-14), and this makes sense by (1-11). Concerning $|D^c b_E|^+$, we just note that if $b_E \in GBV(\mathbb{R}^{n-1})$ — and thus $|D^c b_E|$ is well-defined — then we have

$$|D^c b_E|^+ = |D^c b_E| \llcorner \{v^\wedge > 0\} \quad \text{on Borel sets of } \mathbb{R}^{n-1}.$$

Starting from Theorem 1.7, we can prove a formula for the perimeter of E in terms of v and b_E (see Corollary 3.3) that in turn leads to the following characterization of equality cases in Steiner's inequality in terms of barycenter functions. We recall that, here and in the following results, the assumption $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\})$ is equivalent to asking that $F[v]$ be of finite perimeter, and is thus necessary to make sense of the rigidity problem. In addition we recall that $X \subset \mathbb{R}^m$ is a concentration set for a Borel measure μ on $\mathbb{R}^m$ if $\mu(\mathbb{R}^m \setminus X) = 0$.

Theorem 1.9. *Let $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, and let E be a v-distributed set of finite perimeter. Then, $E \in \mathcal{M}(v)$ if and only if*

$$E_z \text{ is } \mathcal{H}^1\text{-equivalent to a segment for } \mathcal{H}^{n-1}\text{-a.e. } z \in \mathbb{R}^{n-1}, \tag{1-15}$$

$$\nabla b_E(z) = 0 \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } z \in \mathbb{R}^{n-1}, \tag{1-16}$$

$$2[b_E] \leq [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } \{v^\wedge > 0\}, \quad \text{and} \tag{1-17}$$

$$D^c(\tau_M b_\delta)(G) = \int_{G \cap \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}} f \, d(D^c v) \tag{1-18}$$

for every bounded Borel set $G \subset \mathbb{R}^{n-1}$ and for $\mathcal{H}^1$-a.e. $\delta > 0$ and $M > 0$, where $f : \mathbb{R}^{n-1} \to [-\frac{1}{2}, \frac{1}{2}]$ is a

$$\uparrow\downarrow \tfrac{1}{2}[v](z)$$

$$v^\wedge(z) = 0 \qquad\qquad v^\wedge(z) > 0$$
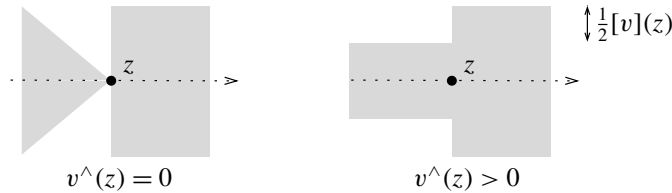
Figure 5. If $E \in \mathcal{M}(v)$, then the jump $[b_E]$ of the barycenter of E can be arbitrarily large on $\{v^\wedge = 0\}$, but is necessarily bounded by half the jump of v on $\{v^\wedge > 0\}$; see (1-17). Moreover, the same rule applies to the Cantorian "jumps", see (1-18) and (1-19).

Borel function; see Figure 5. In particular, $E \in \mathcal{M}(v)$ implies that

$$2|D^c b_E|^+(G) \le |D^c v|(G) \quad \textit{for every Borel set } G \subset \mathbb{R}^{n-1}, \tag{1-19}$$

and that, if K is a concentration set for $D^c v$ and G is a Borel subset of $\{v^\wedge > 0\}$, then

$$\int_\mathbb{R} \mathcal{H}^{n-2}(G \cap \partial^e\{b_E > t\})\, dt = \int_{G \cap S_{b_E} \cap S_v} [b_E]\, d\mathcal{H}^{n-2} + |D^c b_E|^+(G \cap K). \tag{1-20}$$

Remark 1.10. By Theorem 1.7, (1-15) allows us to make sense of ∇b_E, $|D^c b_E|^+$, and $D^c(\tau_M b_\delta)$ (for a.e. $\delta > 0$), and thus to formulate (1-16), (1-18), (1-19), and (1-20). In particular, (1-20) is an immediate consequence of (1-12), (1-16), (1-17), and (1-19).

Theorem 1.9 is a powerful tool in the study of rigidity of equality cases. Indeed, rigidity amounts to asking that b_E be constant on $\{v > 0\}$. Now, b_E is nonconstant (modulo $\mathcal{H}^{n-1}$) on $\{v > 0\}$ if and only if there exists $I \subset \mathbb{R}$ with $\mathcal{H}^1(I) > 0$ such that, if $t \in I$, then $\{\{b_E > t\}, \{b_E \le t\}\}$ is a nontrivial Borel partition of $\{v > 0\}$ (modulo $\mathcal{H}^{n-1}$). In other words, the failure of rigidity is equivalent to saying that $\partial^e\{b_E > t\}$ essentially disconnects $\{v > 0\}$ for every $t \in I$ with $\mathcal{H}^1(I) > 0$. By combining this point of view with the coarea formula (1-20) and with the definition of essential connectedness, we quite easily deduce the following sufficient condition for rigidity.

Theorem 1.11. *If $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$, $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, and the Cantor part $D^c v$ of Dv is concentrated on a Borel set K such that*

$$\{v^\wedge = 0\} \cup S_v \cup K \ \textit{does not essentially disconnect } \{v > 0\}, \tag{1-21}$$

then for every $E \in \mathcal{M}(v)$ there exists $t \in \mathbb{R}$ such that $\mathcal{H}^n\big(E \Delta (t e_n + F[v])\big) = 0$.

Remark 1.12. Note that Theorem 1.11 provides a sufficient condition for rigidity without a priori structural assumption on $F[v]$. In particular, the theorem admits for nontrivial vertical boundaries and vanishing sections, which are excluded in Theorem B by (1-4) and (1-5). (In fact, as shown in Appendix A, Theorem B can be deduced from Theorem 1.11.) We also note that condition (1-21) is clearly not necessary for rigidity as soon as vertical boundaries are present; see Figure 2.

A natural question about equality cases of Steiner's inequality that is left open by Theorem 1.9 is to describe the situation when every $E \in \mathcal{M}(v)$ is obtained by at most countably many vertical translations of parts of $F[v]$. In other words, we want to understand when to expect every $E \in \mathcal{M}(v)$ to satisfy

$$E =_{\mathcal{H}^n} \bigcup_{h \in I} \big(c_h e_n + (F[v] \cap (G_h \times \mathbb{R}))\big), \tag{1-22}$$

where I is at most countable, $\{c_h\}_{h \in I} \subset \mathbb{R}$, and $\{G_h\}_{h \in I}$ is a Borel partition modulo $\mathcal{H}^{n-1}$ of $\{v > 0\}$.

The following theorem shows that this happens when v is of special bounded variation with locally finite jump set. The notion of v-admissible partition of $\{v > 0\}$ used in the theorem is introduced in Definition 1.25; see Section 1E.

Theorem 1.13. *Let $v \in SBV(\mathbb{R}^{n-1}; [0, \infty))$. Assume that $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, and that*

$$S_v \cap \{v^\wedge > 0\} \text{ is locally } \mathcal{H}^{n-2}\text{-finite.} \tag{1-23}$$

Let E be a v-distributed set of finite perimeter. Then, $E \in \mathcal{M}(v)$ if and only if E satisfies (1-22) for a v-admissible partition $\{G_h\}_{h \in I}$ of $\{v > 0\}$ and $2[b_E] \le [v]$ $\mathcal{H}^{n-2}$-a.e. on $\{v^\wedge > 0\}$. Moreover, if these hold, then $|D^c b_E|^+ = 0$.

Remark 1.14. Let us recall that, by definition, $v \in SBV(\mathbb{R}^{n-1})$ if $v \in BV(\mathbb{R}^{n-1})$ and $D^c v = 0$. The approximate discontinuity set S_v of a generic $v \in SBV(\mathbb{R}^{n-1})$ is always countably $\mathcal{H}^{n-2}$-rectifiable, but it may fail to be locally $\mathcal{H}^{n-2}$-finite. If $v \in SBV(\mathbb{R}^{n-1})$ but (1-23) fails, then it may happen that (1-22) does not hold for some $E \in \mathcal{M}(v)$; see Remark 1.32 below.

We close our analysis of equality cases with the following proposition, which shows a general way of producing equality cases in Steiner's inequality that (potentially) do not satisfy the basic structure condition (1-22).

Proposition 1.15. *If $v = v_1 + v_2$, where $v_1, v_2 \in BV(\mathbb{R}^{n-1}; [0, \infty))$, $Dv_1 = D^a v_1$, v_2 is not constant (modulo $\mathcal{H}^{n-1}$) on $\{v > 0\}$, $Dv_2 = D^s v_2$, and $0 < \mathcal{H}^{n-1}(\{v > 0\}) < \infty$, then rigidity fails for v. Indeed, if we set*

$$E = \{x \in \mathbb{R}^n : -\lambda v_2(\mathbf{p}x) - \tfrac{1}{2}v_1(\mathbf{p}x) \le \mathbf{q}x \le \tfrac{1}{2}v_1(\mathbf{p}x) + (1 - \lambda)v_2(\mathbf{p}x)\} \tag{1-24}$$

for $\lambda \in [0, 1] \setminus \{\tfrac{1}{2}\}$, then $E \in \mathcal{M}(v)$ but $\mathcal{H}^n\big(E \Delta (t e_n + F[v])\big) > 0$ for every $t \in \mathbb{R}$. (Note that in (1-24) the choice $\lambda = \tfrac{1}{2}$ gives $E = F[v]$.)

1E. *Characterizations of rigidity.* We now start to discuss the problem of characterizing rigidity of equality cases. We shall analyze this question under different geometric assumptions on the considered Steiner symmetral, and see how different structural assumptions lead to different characterizations.

We begin our analysis by working under the assumption that no vertical boundaries are present where the slice length function v is essentially positive, that is, on $\{v^\wedge > 0\}$. It turns out that, in this case, the sufficient condition (1-21) takes the form

$$\{v^\wedge = 0\} \text{ does not essentially disconnect } \{v > 0\}, \tag{1-25}$$

and that, in turn, this same condition is also necessary to rigidity. Moreover, an alternative characterization can obtained by merely requiring that $F[v]$ be indecomposable.

Theorem 1.16. *Let* $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ *with* $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$ *and*

$$D^s v \llcorner \{v^\wedge > 0\} = 0. \tag{1-26}$$

Then the following are equivalent:

(i) *If* $E \in \mathcal{M}(v)$ *then* $\mathcal{H}^n\big(E \Delta (te_n + F[v])\big) = 0$ *for some* $t \in \mathbb{R}$.

(ii) $\{v^\wedge = 0\}$ *does not essentially disconnect* $\{v > 0\}$.

(iii) $F[v]$ *is indecomposable.*

Remark 1.17. Note that condition (1-26) does not prevent $\partial^* F[v]$ from containing vertical parts, provided they are concentrated where the lower approximate limit of v vanishes. Indeed, (1-26) implies that $D^c v = 0$ (see step one in the proof of Theorem 1.16 in Section 4E), and that S_v is contained in $\{v^\wedge = 0\}$ modulo $\mathcal{H}^{n-2}$. We also note that the equivalence between conditions (ii) and (iii) is actually true whenever $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$; in other words, (1-26) plays no role in proving this equivalence. This is proved in Section 4D, Theorem 4.3.

The situation becomes much more complex when we allow $\partial^* F[v]$ to have vertical parts above $\{v^\wedge > 0\}$. As already noted, simple polyhedral examples, like the one depicted in Figure 2, show that condition (1-21) is not even a viable candidate as a characterization of rigidity in this case. We shall begin our discussion of this problem by solving it in the case of polyhedra and, in fact, in the much broader class of sets introduced in the next definition.

Definition 1.18. Let $v : \mathbb{R}^{n-1} \to [0, \infty)$. We say that $F[v]$ is a *generalized polyhedron* if there exists a *finite* disjoint family of indecomposable sets of finite perimeter and volume $\{A_j\}_{j \in J}$ in $\mathbb{R}^{n-1}$, and a family of functions $\{v_j\}_{j \in J} \subset W^{1,1}(\mathbb{R}^{n-1})$, such that

$$v = \sum_{j \in J} v_j 1_{A_j}, \tag{1-27}$$

$$\big(\{v^\wedge = 0\} \setminus \{v = 0\}^{(1)}\big) \cup S_v \subset_{\mathcal{H}^{n-2}} \bigcup_{j \in J} \partial^e A_j. \tag{1-28}$$

(Here and in the following, $A \subset_{\mathcal{H}^k} B$ stands for $\mathcal{H}^k(A \setminus B) = 0$.)

Remark 1.19. Condition (1-28) amounts to requiring that v can jump or essentially vanish on $\{v > 0\}$ only inside the essential boundaries of the sets A_j. For example, if $\{A_j\}_{j \in J}$ is a finite disjoint family of bounded open sets with Lipschitz boundary in $\mathbb{R}^{n-1}$, $\{v_j\}_{j \in J} \subset C^1(\mathbb{R}^{n-1})$, and $v_j > 0$ on A_j for every $j \in J$, then $v = \sum_{j \in J} v_j 1_{A_j}$ defines a generalized polyhedron $F[v]$. Note that in this case (1-28) holds, since v can jump only over the boundaries of the A_j, so that $S_v \subset \bigcup_{j \in J} \partial A_j$, while $\{v_j = 0\} \cap \overline{A_j} \subset \partial A_j$ for every $j \in J$.

Theorem 1.20. *If* $v : \mathbb{R}^{n-1} \to [0, \infty)$ *is such that* $F[v]$ *is a generalized polyhedron, then the following two statements are equivalent*:

(i) *If $E \in \mathcal{M}(v)$ then $\mathcal{H}^n\big(E \Delta (te_n + F[v])\big) = 0$ for some $t \in \mathbb{R}$.*

(ii) *For every $\varepsilon > 0$ the set $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ does not essentially disconnect $\{v > 0\}$.*

Remark 1.21. In the example depicted in Figure 2, the set $\{v^\wedge = 0\} \cap \{v > 0\}^{(1)}$ is empty, the set $\{[v] > 0\}$ essentially disconnects $\{v > 0\}$, but there is no $\varepsilon > 0$ such that $\{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$. Indeed, in this case, rigidity holds.

Note that, if $F[v]$ is a generalized polyhedron, then $v \in SBV(\mathbb{R}^{n-1})$ with S_v locally $\mathcal{H}^{n-2}$-rectifiable, so that v satisfies the assumptions of Theorem 1.13. We now discuss the rigidity problem in this more general situation.

As shown by Example 1.22 below, condition (ii) in Theorem 1.20 is not even a sufficient condition for rigidity under the assumptions on v considered in Theorem 1.13. A key remark here is that, in the situations considered in Theorem 1.16 and Theorem 1.20, we can create failure of rigidity by performing a vertical translation of $F[v]$ above a single part of $\{v > 0\}$. For example, when condition (ii) in Theorem 1.20 fails, there exist $\varepsilon > 0$ and a nontrivial Borel partition $\{G_+, G_-\}$ of $\{v > 0\}$ modulo $\mathcal{H}^{n-1}$ such that

$$\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\} \cup \{[v] > \varepsilon\}.$$

In that case, as we shall prove later on, the v-distributed set $E(t)$ defined as

$$E(t) = \big((te_n + F[v]) \cap (G_+ \times \mathbb{R})\big) \cup (F[v] \cap (G_- \times \mathbb{R})), \quad t \in \mathbb{R},$$

and obtained by a single vertical translation of $F[v]$ above G_+, satisfies $P(E(t)) = P(F[v])$ whenever $t \in (0, \varepsilon/2)$. (Moreover, when condition (1-25) fails, we have $E(t) \in \mathcal{M}(v)$ for every $t \in \mathbb{R}$.) However, there may be situations in which violating rigidity by a single vertical translation of $F[v]$ is impossible, but where this task can be accomplished by simultaneously performing countably many independent vertical translations of $F[v]$. An example is obtained as follows.

Example 1.22. We construct a function $v : \mathbb{R}^2 \to [0, \infty)$ in such a way that $v \in SBV(\mathbb{R}^2)$, S_v is locally $\mathcal{H}^1$-rectifiable, the set $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ does not essentially disconnect $\{v > 0\}$ for any $\varepsilon > 0$, but, nevertheless, rigidity fails. Given $t \in \mathbb{R}$ and $\ell > 0$, denote by $Q(t, \ell)$ the open square in $\mathbb{R}^2$ with center at $(t, 0)$, sides parallel to the direction $(1, 1)$ and $(1, -1)$, and diagonal of length 2ℓ. Then we set $u_1 = 1_{Q(0,1)}$, and define a sequence $\{u_j\}_{j \in \mathbb{N}}$ of piecewise constant functions

$$u_2 = u_1 - \tfrac{1}{2} 1_{Q(-3/4, 1/4)} + \tfrac{1}{2} 1_{Q(3/4, 1/4)},$$
$$u_3 = u_2 - \tfrac{1}{4} 1_{Q(-15/16, 1/16)} + \tfrac{1}{4} 1_{Q(-9/16, 1/16)} - \tfrac{1}{4} 1_{Q(9/16, 1/16)} + \tfrac{1}{4} 1_{Q(15/16, 1/16)},$$

etc.; see Figure 6. This sequence has pointwise limit $v \in SBV(\mathbb{R}^2; [0, \infty))$ such that $\{v > 0\} = Q(0, 1)$ and $Dv = D^s v$. In particular, if we define E as in (1-24) with $\lambda = 0$, $v_1 = 0$, and $v_2 = v$, then, by Proposition 1.15, $E \in \mathcal{M}(v)$. Since $b_E = \tfrac{1}{2} v$, we easily see that (1-34), and thus (1-22), holds; in other words, E is obtained by countably many vertical translations of $F[v]$ over suitable disjoint Borel sets G_h, $h \in \mathbb{N}$. At the same time, any set E_0 obtained by a vertical translation of $F[v]$ over one (or over finitely many) of the G_h is bound to violate the necessary condition for equality, $2[b_{E_0}] \leq [v]$ $\mathcal{H}^{n-2}$-a.e. on $S_v \cap \{v^\wedge > 0\}$, as the infimum of $[v]$ on $\partial^e G_h \cap S_v \cap \{v^\wedge > 0\}$ is zero for every $h \in \mathbb{N}$. We also note
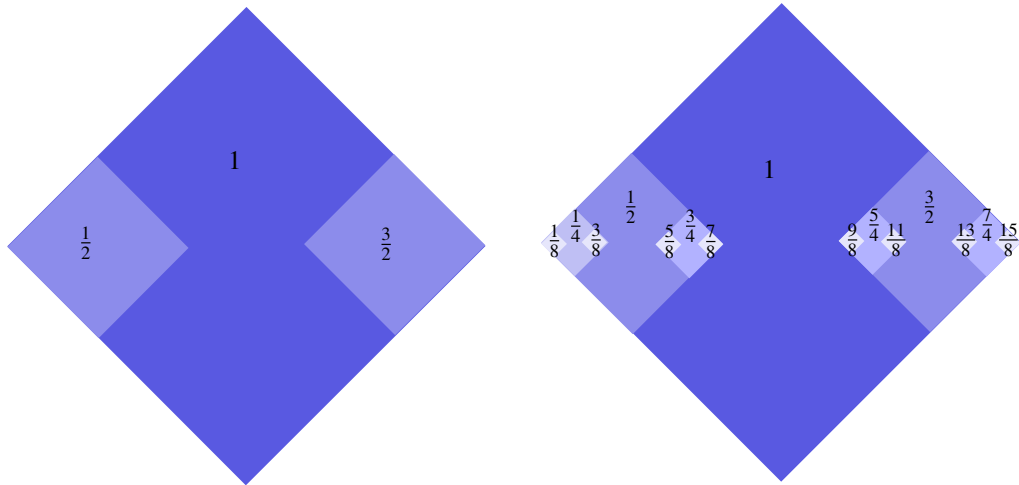
Figure 6. The functions u_2 and u_4 in the construction of Example 1.22.

that, as a simple computation shows, $S_v \cap \{v^\wedge > 0\}$ is not only countably $\mathscr{H}^1$-rectifiable in $\mathbb{R}^2$ but actually $\mathscr{H}^1$-finite (thus, it is locally $\mathscr{H}^1$-rectifiable).

All the above considerations finally suggest the following condition, which, in turn, characterizes rigidity under the assumptions on v considered in Theorem 1.13. We begin by recalling the definition of a Caccioppoli partition.

Definition 1.23. Let $G \subset \mathbb{R}^{n-1}$ be a set of finite perimeter and let $\{G_h\}_{h \in I}$ be an at most countable Borel partition of G modulo $\mathscr{H}^{n-1}$. (That is, I is a finite or countable set with $\#I \geq 2$, $G =_{\mathscr{H}^{n-1}} \bigcup_{h \in I} G_h$, $\mathscr{H}^{n-1}(G_h) > 0$ for every $h \in I$, and $\mathscr{H}^{n-1}(G_h \cap G_k) = 0$ for every $h, k \in I$, $h \neq k$.) We say that $\{G_h\}_{h \in I}$ is a *Caccioppoli partition of* G if $\sum_{h \in I} P(G_h) < \infty$.

Remark 1.24. When G is an open set and $\{G_h\}_{h \in I}$ is an at most countable Borel partition of G modulo $\mathscr{H}^{n-1}$, then, according to [Ambrosio et al. 2000, Definition 4.16], $\{G_h\}_{h \in I}$ is a Caccioppoli partition of G if $\sum_{h \in I} P(G_h; G) < \infty$. Of course, if we assume in addition that G is of finite perimeter, then $\sum_{h \in I} P(G_h; G) < \infty$ is equivalent to $\sum_{h \in I} P(G_h) < \infty$. Thus Definition 1.23 and [Ambrosio et al. 2000, Definition 4.16] agree in their common domain of applicability (that is, on open sets of finite perimeter).

Definition 1.25. Let $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$, and let $\{G_h\}_{h \in I}$ be an at most countable Borel partition of $\{v > 0\}$. We say that $\{G_h\}_{h \in I}$ is a *v-admissible partition* of $\{v > 0\}$ if $\{G_h \cap B_R \cap \{v > \delta\}\}_{h \in I}$ is a Caccioppoli partition of $\{v > \delta\} \cap B_R$ for every $\delta > 0$ such that $\{v > \delta\}$ is of finite perimeter and for every $R > 0$.

Definition 1.26. One says that $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ satisfies the *mismatched stairway property* if the following holds: If $\{G_h\}_{h \in I}$ is a v-admissible partition of $\{v > 0\}$ and if $\{c_h\}_{h \in I} \subset \mathbb{R}$ is a sequence with

$c_h \neq c_k$ whenever $h \neq k$, then there exist $h_0, k_0 \in I$ with $h_0 \neq k_0$ and a Borel set Σ with

$$\Sigma \subset \partial^e G_{h_0} \cap \partial^e G_{k_0} \cap \{v^\wedge > 0\}, \quad \mathcal{H}^{n-2}(\Sigma) > 0, \tag{1-29}$$

such that

$$[v](z) < 2|c_{h_0} - c_{k_0}| \quad \text{for all } z \in \Sigma. \tag{1-30}$$

Remark 1.27. The terminology adopted here intends to suggest the following idea. One considers a v-admissible partition $\{G_h\}_{h \in I}$ of $\{v > 0\}$ such that $\{v > 0\}^{(1)} \cap \bigcup_{h \in I} \partial^e G_h$ is contained in $\{v^\wedge = 0\} \cup S_v$. Next, one modifies $F[v]$ by performing vertical translations c_h above each G_h, thus constructing a new set E having a "stairway-like" barycenter function. This new set will have the same perimeter of $F[v]$, and thus will violate rigidity if $\#I \geq 2$, provided *all* the steps of the stairway match the jumps of v, in the sense that $2[b_E] = 2|c_h - c_k| \leq [v]$ on each $\partial^e G_h \cap \partial^e G_k \cap \{v^\wedge > 0\}$. Thus, when all equality cases have a stairway-like barycenter function, we expect rigidity to be equivalent to asking that every such stairway has *at least one* step that is *mismatched* with respect to $[v]$; compare with (1-30).

Remark 1.28. If $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ has the mismatched stairway property, then, for every $\varepsilon > 0$, $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ does not essentially disconnect $\{v > 0\}$. In particular, $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$, $\{v > 0\}$ is essentially connected, and although it may still happen that $\{v^\wedge = 0\} \cup S_v$ essentially disconnects $\{v > 0\}$, in this case one has

$$\mathcal{H}^{n-2}\text{-essinf}_{S_v \cap \{v^\wedge > 0\}}[v] = 0.$$

We prove the claim arguing by contradiction. If $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$, then there exist $\varepsilon > 0$ and a nontrivial Borel partition $\{G_+, G_-\}$ of $\{v > 0\}$ modulo $\mathcal{H}^{n-1}$ such that $\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$. Since (2-9) below implies $\{v^\wedge > 0\} \subset \{v > 0\}^{(1)}$, we have

$$\{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{[v] > \varepsilon\}, \tag{1-31}$$

so that, for every $\delta > 0$ (and since $\{v > \delta\}^{(1)} \cap \partial^e G_+ = \{v > \delta\}^{(1)} \cap \partial^e G_-$),

$$\{v > \delta\}^{(1)} \cap \partial^e G_+ = \{v > \delta\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{[v] > \varepsilon\}. \tag{1-32}$$

If we set $G_{\pm\delta} = G_\pm \cap \{v > \delta\}$, then $\partial^e G_{\pm\delta} \subset \partial^e\{v > \delta\} \cup (\{v > \delta\}^{(1)} \cap \partial^e G_\pm)$, and, by (1-32), $\partial^e G_{\pm\delta} \subset_{\mathcal{H}^{n-2}} \partial^e\{v > \delta\} \cup \{[v] > \varepsilon\}$. Since $[v] \in L^1(\mathcal{H}^{n-2} \llcorner S_v)$, we find $\mathcal{H}^{n-2}(\{[v] > t\}) < \infty$ for every $t > 0$, and, in particular

$$P(G_{+\delta}) + P(G_{-\delta}) \leq 2P(\{v > \delta\}) + 2\mathcal{H}^{n-2}(\{[v] > \varepsilon\}) < \infty$$

whenever $\{v > \delta\}$ is of finite perimeter. This shows that $\{G_+, G_-\}$ is a v-admissible partition. If we now set $I = \{+, -\}$, $c_+ = \varepsilon/2$, and $c_- = 0$, then I, $\{G_h\}_{h \in I}$, and $\{c_h\}_{h \in I}$ are admissible in the mismatched stairway property. By the mismatched stairway property, there exists a Borel set $\Sigma \subset \{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_-$ such that $[v] < 2|c_+ - c_-| = \varepsilon$ on Σ and $\mathcal{H}^{n-2}(\Sigma) > 0$, a contradiction to (1-31).

It turns out that if v is a *SBV*-function with locally finite jump set, then rigidity is characterized by the mismatched stairway property.

Theorem 1.29. *If* $v \in SBV(\mathbb{R}^{n-1}; [0, \infty))$, $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, *and* $S_v \cap \{v^\wedge > 0\}$ *is locally* $\mathcal{H}^{n-2}$-*finite, then the following two statements are equivalent*:

(i) *If* $E \in \mathcal{M}(v)$, *then* $\mathcal{H}^n\big(E \Delta (te_n + F[v])\big) = 0$ *for some* $t \in \mathbb{R}$.

(ii) v *has the mismatched stairway property.*

The question of a geometric characterization of rigidity when $v \in BV$ is thus left open. The considerable complexity of the mismatched stairway property may be seen as a negative indication about the tractability of this problem. In the planar case, due to the trivial topology of the real line, these difficulties can be overcome, and we obtain the following complete result.

Theorem 1.30. *If* $v \in BV(\mathbb{R}; [0, \infty))$ *and* $\mathcal{H}^1(\{v > 0\}) < \infty$, *then the following are equivalent*:

(i) *If* $E \in \mathcal{M}(v)$, *then* $\mathcal{H}^2\big(E \Delta (te_2 + F[v])\big) = 0$ *for some* $t \in \mathbb{R}$.

(ii) $\{v > 0\}$ *is* $\mathcal{H}^1$-*equivalent to a bounded open interval* (a, b), $v \in W^{1,1}(a, b)$, *and* $v^\wedge > 0$ *on* (a, b).

(iii) $F[v]$ *is an indecomposable set that has no vertical boundary above* $\{v^\wedge > 0\}$, *i.e.*,

$$\mathcal{H}^1\big(\{x \in \partial^* F[v] : \boldsymbol{q}\nu_{F[v]}(x) = 0, v^\wedge(\boldsymbol{p}x) > 0\}\big) = 0. \tag{1-33}$$

The extension of our results to the case of the localized Steiner inequality is discussed in Appendix A. In particular, we shall explain how to derive Theorem B from Theorem 1.11 via an approximation argument.

1F. *Some closing remarks.* We conclude this introduction with a few remarks of more technical nature. The first two remarks deal with the issue addressed in Theorem 1.13, namely, understanding when equality cases are necessarily obtained by countably many vertical translations of the Steiner symmetral; see (1-22). Theorem 1.13 ensures this is the case if $v \in SBV(\mathbb{R}^{n-1})$ with $S_v \cap \{v^\wedge > 0\}$ locally $\mathcal{H}^{n-2}$-finite. In the following two remarks we show that, if we merely assume that $v \in SBV(\mathbb{R}^{n-1})$, then we can indeed construct equality cases that do not satisfy (1-22).

Remark 1.31. Condition (1-22) can be reformulated in terms of a property of the barycenter function. Indeed, (1-22) is equivalent to asking that

$$b_E = \sum_{h \in I} c_h 1_{G_h} \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1} \tag{1-34}$$

for I, $\{c_h\}_{h \in I}$ and $\{G_h\}_{h \in I}$ as in (1-22). It should be noted that, *if no additional conditions are assumed on the partition* $\{G_h\}_{h \in I}$, then (1-34) is *not* equivalent to saying that b_E has "countable range". An example is obtained as follows. Let K be the middle-third Cantor set in $[0, 1]$, let $\{G_h\}_{h \in \mathbb{N}}$ be the disjoint family of open intervals such that $K = [0, 1] \setminus \bigcup_{h \in \mathbb{N}} G_h$, and let $\{c_h\}_{h \in \mathbb{N}} \subset \mathbb{R}$ be such that the Cantor function u_K satisfies $u_K = c_h$ on G_h. In this way, $u_K = \sum_{h \in \mathbb{N}} c_h 1_{G_h}$ on $[0, 1] \setminus K$, thus $\mathcal{H}^1$-a.e. on $[0, 1]$. Of course, since u_K is a nonconstant, continuous, and increasing function, it does not have "countable range" in any reasonable sense. At the same time, if we set $v(z) = 1_{[0,1]}(z) \operatorname{dist}(z, K)$ for $z \in \mathbb{R}$, then v is a Lipschitz function on $\mathbb{R}$ (thus it satisfies all the assumptions in Theorem 1.13) and the set

$$E = \{x \in \mathbb{R}^2 : u_K(\boldsymbol{p}x) - \tfrac{1}{2}v(\boldsymbol{p}x) < \boldsymbol{q}x < u_K(\boldsymbol{p}x) + \tfrac{1}{2}v(\boldsymbol{p}x)\}$$

is such that $E \in \mathcal{M}(v)$, as one can check by Corollary 3.3 and Corollary 3.4 in Section 3B. We also note that, in this example, $|D^c b_E| \llcorner \{v^\wedge = 0\} \neq 0$, while $|D^c b_E|^+ = 0$.

Remark 1.32. We now describe the example introduced in Remark 1.14. Given $\{q_h\}_{h \in \mathbb{N}} = \mathbb{Q} \cap [0, 1]$ and $\{\alpha_h\}_{h \in \mathbb{N}} \in (0, \infty)$ such that $\sum_{h \in \mathbb{N}} \alpha_h < \infty$, we can define $v \in SBV(\mathbb{R})$ such that $\mathcal{H}^1(\{v > 0\}) = 1$ and $Dv = D^s v = D^j v$, by setting

$$v(t) = \sum_{\{h \in \mathbb{N}: q_h < t \leq 1\}} \alpha_h = \sum_{h \in \mathbb{N}} \alpha_h 1_{(q_h, 1]}(t), \quad t \in \mathbb{R}.$$

If we let $v_1 = 0$, $v_2 = v$, and, say, $\lambda = 0$, in Proposition 1.15 below, then we obtain a set $E \in \mathcal{M}(v)$. At the same time, (1-34), and thus (1-22), cannot hold, as $b_E = \frac{1}{2} v \; \mathcal{H}^1$-a.e. on $\mathbb{R}$ and v is *strictly* increasing on $[0, 1]$. (The requirement that the sets G_h in (1-34) are mutually disjoint modulo $\mathcal{H}^{n-1}$ plays a crucial role in here, of course.) Note that, as expected, $S_v \cap \{v^\wedge > 0\} = \mathbb{Q} \cap [0, 1]$ is not locally $\mathcal{H}^0$-finite.

The following final remark is instead concerned with the characterization presented in Theorem 1.29 in terms of the mismatched stairway property.

Remark 1.33. Is it important to observe that, in order to characterize rigidity, only v-admissible partitions of $\{v > 0\}$ have to be considered in the definition of the mismatched stairway property. Indeed, let $n = 2$ and set $v = 1_{(0,1)} \in SBV(\mathbb{R}; [0, \infty))$, so that rigidity holds for v. Now let $\{G_h\}_{h \in \mathbb{N}}$ be the family of open intervals used to define the middle-third Cantor set K, so that $K = [0, 1] \setminus \bigcup_{h \in \mathbb{N}} G_h$. Note that $\{G_h\}_{h \in \mathbb{N}}$ is a nontrivial countable Borel partition of $\{v > 0\} = (0, 1)$ modulo $\mathcal{H}^1$. However, since $\partial^e G_h \cap \partial^e G_k = \varnothing$ whenever $h \neq k$, it is not possible to find a set Σ satisfying (1-29), whatever choice of $\{c_h\}_{h \in \mathbb{N}}$ we make. In particular, if we did not restrict the partitions in Definition 1.26 to v-admissible partitions, then this particular v (satisfying rigidity) would not have the mismatched stairway property. Note of course that, in this example, $\sum_{h \in \mathbb{N}} P(G_h \cap \{v > \delta\} \cap B_R) = \infty$ for every $\delta, R > 0$.

2. Notions from geometric measure theory

We gather here some notions from geometric measure theory needed in the sequel, referring to [Ambrosio et al. 2000; Maggi 2012] for further details. We start by reviewing our general notation in $\mathbb{R}^n$. We denote by $B(x, r)$ the open Euclidean ball of radius $r > 0$ and center $x \in \mathbb{R}^n$. Given $x \in \mathbb{R}^n$ and $v \in S^{n-1}$ we denote by $H_{x,v}^+$ and $H_{x,v}^-$ the complementary half-spaces

$$H_{x,v}^+ = \{y \in \mathbb{R}^n : (y - x) \cdot v \geq 0\}, \quad H_{x,v}^- = \{y \in \mathbb{R}^n : (y - x) \cdot v \leq 0\}. \tag{2-1}$$

Finally, we decompose $\mathbb{R}^n$ as the product $\mathbb{R}^{n-1} \times \mathbb{R}$, and denote by $\boldsymbol{p} : \mathbb{R}^n \to \mathbb{R}^{n-1}$ and $\boldsymbol{q} : \mathbb{R}^n \to \mathbb{R}$ the corresponding horizontal and vertical projections, so that

$$x = (\boldsymbol{p}x, \boldsymbol{q}x) = (x', x_n), \quad x' = (x_1, \ldots, x_{n-1}) \quad \text{for all } x \in \mathbb{R}^n,$$

and define the vertical cylinder of center $x \in \mathbb{R}^n$ and radius $r > 0$, and the $(n-1)$-dimensional ball in $\mathbb{R}^{n-1}$ of center $z \in \mathbb{R}^{n-1}$ and radius $r > 0$ by setting, respectively,

$$\boldsymbol{C}_{x,r} = \{y \in \mathbb{R}^n : |\boldsymbol{p}x - \boldsymbol{p}y| < r, |\boldsymbol{q}x - \boldsymbol{q}y| < r\}, \quad \boldsymbol{D}_{z,r} = \{w \in \mathbb{R}^{n-1} : |w - z| < r\}.$$

In this way, $C_{x,r} = D_{px,r} \times (qx - r, qx + r)$. We shall use the following two notions of convergence for Lebesgue measurable subsets of $\mathbb{R}^n$. Given Lebesgue measurable sets $\{E_h\}_{h \in \mathbb{N}}$ and E in $\mathbb{R}^n$, we shall say that E_h locally converge to E, and write

$$E_h \xrightarrow{\text{loc}} E \quad \text{as } h \to \infty,$$

provided $\mathcal{H}^n((E_h \Delta E) \cap K) \to 0$ as $h \to \infty$ for every compact set $K \subset \mathbb{R}^n$; we say that E_h converge to E as $h \to \infty$, and write $E_h \to E$, provided $\mathcal{H}^n(E_h \Delta E) \to 0$ as $h \to \infty$.

2A. *Density points and approximate limits.* If E is a Lebesgue measurable set in $\mathbb{R}^n$ and $x \in \mathbb{R}^n$, then we define the *upper and lower n-dimensional densities* of E at x as

$$\theta^*(E, x) = \limsup_{r \to 0^+} \frac{\mathcal{H}^n(E \cap B(x, r))}{\omega_n r^n} \quad \text{and} \quad \theta_*(E, x) = \liminf_{r \to 0^+} \frac{\mathcal{H}^n(E \cap B(x, r))}{\omega_n r^n}$$

respectively. In this way we define two Borel functions on $\mathbb{R}^n$ that agree a.e. on $\mathbb{R}^n$. In particular, the *n-dimensional density* of E at x,

$$\theta(E, x) = \lim_{r \to 0^+} \frac{\mathcal{H}^n(E \cap B(x, r))}{\omega_n r^n},$$

is defined for a.e. $x \in \mathbb{R}^n$, and $\theta(E, \cdot)$ is a Borel function on $\mathbb{R}^n$ (up to extending it by a constant value on the $\mathcal{H}^n$-negligible set $\{\theta^*(E, \cdot) > \theta_*(E, \cdot)\}$). Correspondingly, for $t \in [0, 1]$, we define

$$E^{(t)} = \{x \in \mathbb{R}^n : \theta(E, x) = t\}. \tag{2-2}$$

By the Lebesgue differentiation theorem, $\{E^{(0)}, E^{(1)}\}$ is a partition of $\mathbb{R}^n$ up to an $\mathcal{H}^n$-negligible set. It is useful to keep in mind that

$$x \in E^{(1)} \quad \text{if and only if} \quad E_{x,r} \xrightarrow{\text{loc}} \mathbb{R}^n \text{ as } r \to 0^+,$$

$$x \in E^{(0)} \quad \text{if and only if} \quad E_{x,r} \xrightarrow{\text{loc}} \varnothing \text{ as } r \to 0^+,$$

where $E_{x,r}$ denotes the blow-up of E at x at scale r, defined as

$$E_{x,r} = \frac{E - x}{r} = \left\{ \frac{y - x}{r} : y \in E \right\}, \quad x \in \mathbb{R}^n, r > 0.$$

The set $\partial^e E = \mathbb{R}^n \setminus (E^{(0)} \cup E^{(1)})$ is called the *essential boundary* of E. Thus, in general, we only have $\mathcal{H}^n(\partial^e E) = 0$, but we do not know $\partial^e E$ to be "$(n-1)$-dimensional" in any sense. Strictly related to the notion of density is that of approximate upper and lower limits of a measurable function. Given a Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ we define the (weak) *approximate upper and lower limits* of f at $x \in \mathbb{R}^n$ as

$$f^\vee(x) = \inf\{t \in \mathbb{R} : \theta(\{f > t\}, x) = 0\} = \inf\{t \in \mathbb{R} : \theta(\{f < t\}, x) = 1\},$$

$$f^\wedge(x) = \sup\{t \in \mathbb{R} : \theta(\{f < t\}, x) = 0\} = \sup\{t \in \mathbb{R} : \theta(\{f > t\}, x) = 1\}.$$

As it turns out, $f^\vee$ and $f^\wedge$ are Borel functions with values on $\mathbb{R} \cup \{\pm\infty\}$ defined *at every point x* of $\mathbb{R}^n$, and they do not depend on the Lebesgue representative chosen for the function f. Moreover, for

$\mathcal{H}^n$-a.e. $x \in \mathbb{R}^n$, we have that $f^\vee(x) = f^\wedge(x) \in \mathbb{R} \cup \{\pm\infty\}$, so that the *approximate discontinuity* set of f, $S_f = \{f^\wedge < f^\vee\}$, satisfies $\mathcal{H}^n(S_f) = 0$. On noticing that, though $f^\wedge$ and $f^\vee$ may take infinite values on S_f, the difference $f^\vee(x) - f^\wedge(x)$ is always well-defined in $\mathbb{R} \cup \{\pm\infty\}$ for $x \in S_f$, we define the *approximate jump* of f as the Borel function $[f] : \mathbb{R}^n \to [0, \infty]$ defined by

$$[f](x) = \begin{cases} f^\vee(x) - f^\wedge(x) & \text{if } x \in S_f, \\ 0 & \text{if } x \in \mathbb{R}^n \setminus S_f, \end{cases}$$

so that $S_f = \{[f] > 0\}$. Finally, the *approximate average* of f is the Borel function $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ defined as

$$\tilde{f}(x) = \begin{cases} \frac{1}{2}(f^\vee(x) + f^\wedge(x)) & \text{if } x \in \mathbb{R}^n \setminus \{f^\wedge = -\infty, f^\vee = +\infty\}, \\ 0 & \text{if } x \in \{f^\wedge = -\infty, f^\vee = +\infty\}. \end{cases} \tag{2-3}$$

The motivation behind definition (2-3) is that (in step two of the proof of Theorem 3.1) we want the limit relation

$$\tilde{f}(x) = \lim_{M \to \infty} \widetilde{\tau_M(f)}(x) = \lim_{M \to \infty} \frac{1}{2}(\tau_M(f^\vee) + \tau_M(f^\wedge)) \quad \text{for all } x \in \mathbb{R}^n \tag{2-4}$$

to hold for every Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$, where here and in the rest of the paper we set

$$\tau_M(s) = \max\{-M, \min\{M, s\}\}, \quad s \in \mathbb{R} \cup \{\pm\infty\}. \tag{2-5}$$

The validity of (2-4) is easily checked by noticing that

$$\tau_M(f)^\wedge = \tau_M(f^\wedge), \quad \tau_M(f)^\vee = \tau_M(f^\vee), \quad \widetilde{\tau_M(f)}(x) = \frac{1}{2}\tau_M(f^\vee) + \tau_M(f^\wedge). \tag{2-6}$$

With these definitions at hand, we note the validity of the following properties, which follow easily from the above definitions, and hold for every Lebesgue measurable $f : \mathbb{R}^n \to \mathbb{R}$ and for every $t \in \mathbb{R}$:

$$\{|f|^\vee < t\} = \{-t < f^\wedge\} \cap \{f^\vee < t\}, \tag{2-7}$$

$$\{f^\vee < t\} \subset \{f < t\}^{(1)} \subset \{f^\vee \le t\}, \tag{2-8}$$

$$\{f^\wedge > t\} \subset \{f > t\}^{(1)} \subset \{f^\wedge \ge t\}. \tag{2-9}$$

(Note that all the inclusions may be strict, that we also have $\{f < t\}^{(1)} = \{f^\vee < t\}^{(1)}$, and that all the other analogous relations hold.) Moreover, if $f, g : \mathbb{R}^n \to \mathbb{R}$ are Lebesgue measurable functions and $f = g$ $\mathcal{H}^n$-a.e. on a Borel set E, then

$$f^\vee(x) = g^\vee(x), \quad f^\wedge(x) = g^\wedge(x), \quad [f](x) = [g](x) \quad \text{for all } x \in E^{(1)}. \tag{2-10}$$

If $f : \mathbb{R}^n \to \mathbb{R}$ and $A \subset \mathbb{R}^n$ are Lebesgue measurable, and $x \in \mathbb{R}^n$ is such that $\theta^*(A, x) > 0$, then we say that $t \in \mathbb{R} \cup \{\pm\infty\}$ is the *approximate limit of f at x with respect to A*, and write $t = \operatorname{aplim}(f, A, x)$, if

$$\theta(\{|f - t| > \varepsilon\} \cap A; x) = 0 \quad \text{for all } \varepsilon > 0 \quad (t \in \mathbb{R}),$$

$$\theta(\{f < M\} \cap A; x) = 0 \quad \text{for all } M > 0 \quad (t = +\infty),$$

$$\theta(\{f > -M\} \cap A; x) = 0 \quad \text{for all } M > 0 \quad (t = -\infty).$$

We say that $x \in S_f$ is a *jump point* of f if there exists $\nu \in S^{n-1}$ such that

$$f^{\vee}(x) = \mathrm{aplim}(f, H_{x,\nu}^+, x), \quad f^{\wedge}(x) = \mathrm{aplim}(f, H_{x,\nu}^-, x).$$

If this is the case we set $\nu = \nu_f(x)$, the *approximate jump direction* of f at x. We denote by J_f the set of approximate jump points of f, so that $J_f \subset S_f$; moreover, $\nu_f : J_f \to S^{n-1}$ is a Borel function. It will be particularly useful to keep in mind the following proposition; see [Cagnetti et al. 2013, Proposition 2.2] for a proof.

Proposition 2.1. *We have that $x \in J_f$ if and only if, for every $\tau \in (f^{\wedge}(x), f^{\vee}(x))$,*

$$\{f > \tau\}_{x,r} \xrightarrow{\mathrm{loc}} H_{0,\nu}^+ \quad and \quad \{f < \tau\}_{x,r} \xrightarrow{\mathrm{loc}} H_{0,\nu}^- \quad as \ r \to 0^+. \tag{2-11}$$

Finally, if $f : \mathbb{R}^n \to \mathbb{R}$ is Lebesgue measurable, then we say f is *approximately differentiable* at $x \in S_f^c$ provided $f^{\wedge}(x) = f^{\vee}(x) \in \mathbb{R}$ and there exists $\xi \in \mathbb{R}^n$ such that

$$\mathrm{aplim}(g, \mathbb{R}^n, x) = 0,$$

where $g(y) = (f(y) - \tilde{f}(x) - \xi \cdot (y - x))/|y - x|$ for $y \in \mathbb{R}^n \setminus \{x\}$. If this is the case, then ξ is uniquely determined, we set $\xi = \nabla f(x)$, and call $\nabla f(x)$ the *approximate differential* of f at x. The localization property (2-10) holds also for approximate differentials: precisely, if $f, g : \mathbb{R}^n \to \mathbb{R}$ are Lebesgue measurable functions, $f = g$ $\mathscr{H}^n$-a.e. on a Borel set E, and f is approximately differentiable $\mathscr{H}^n$-a.e. on E, then g is approximately differentiable $\mathscr{H}^n$-a.e. on E too, with

$$\nabla f(x) = \nabla g(x) \quad \text{for } \mathscr{H}^n\text{-a.e. } x \in E. \tag{2-12}$$

2B. *Rectifiable sets and functions of bounded variation.* Let $1 \le k \le n$, $k \in \mathbb{N}$. A Borel set $M \subset \mathbb{R}^n$ is *countably $\mathscr{H}^k$-rectifiable* if there are Lipschitz functions $f_h : \mathbb{R}^k \to \mathbb{R}^n$, $h \in \mathbb{N}$, such that $M \subset_{\mathscr{H}^k} \bigcup_{h \in \mathbb{N}} f_h(\mathbb{R}^k)$. We further say that M is *locally $\mathscr{H}^k$-rectifiable* if $\mathscr{H}^k(M \cap K) < \infty$ for every compact set $K \subset \mathbb{R}^n$, or, equivalently, if $\mathscr{H}^k \llcorner M$ is a Radon measure on $\mathbb{R}^n$. Hence, for a locally $\mathscr{H}^k$-rectifiable set M in $\mathbb{R}^n$ the following definition is well-posed: we say that M has a k-dimensional subspace L of $\mathbb{R}^n$ as its *approximate tangent plane* at $x \in \mathbb{R}^n$, $L = T_x M$, if $\mathscr{H}^k \llcorner (M - x)/r \rightharpoonup \mathscr{H}^k \llcorner L$ as $r \to 0^+$ weakly star in the sense of Radon measures. It turns out that $T_x M$ exists and is uniquely defined at $\mathscr{H}^k$-a.e. $x \in M$. Moreover, given two locally $\mathscr{H}^k$-rectifiable sets M_1 and M_2 in $\mathbb{R}^n$, we have $T_x M_1 = T_x M_2$ for $\mathscr{H}^k$-a.e. $x \in M_1 \cap M_2$.

A Lebesgue measurable set $E \subset \mathbb{R}^n$ is said to be of *locally finite perimeter* in $\mathbb{R}^n$ if there exists an $\mathbb{R}^n$-valued Radon measure μ_E, called the *Gauss–Green measure* of E, such that

$$\int_E \nabla \varphi(x) \, dx = \int_{\mathbb{R}^n} \varphi(x) \, d\mu_E(x) \quad \text{for all } \varphi \in C_c^1(\mathbb{R}^n).$$

The relative perimeter of E in $A \subset \mathbb{R}^n$ is then defined by setting $P(E; A) = |\mu_E|(A)$, while the perimeter of E is $P(E) = P(E; \mathbb{R}^n)$. The *reduced boundary* of E is the set $\partial^* E$ of those $x \in \mathbb{R}^n$ such that

$$\nu_E(x) = \lim_{r \to 0^+} \frac{\mu_E(B(x,r))}{|\mu_E|(B(x,r))} \quad \text{exists and belongs to } S^{n-1}.$$

The Borel function $\nu_E : \partial^* E \to S^{n-1}$ is called the *measure-theoretic outer unit normal* to E. It turns out that $\partial^* E$ is a locally $\mathcal{H}^{n-1}$-rectifiable set in $\mathbb{R}^n$ [Maggi 2012, Corollary 16.1], that $\mu_E = \nu_E \mathcal{H}^{n-1} \llcorner \partial^* E$, and that

$$\int_E \nabla \varphi(x) \, dx = \int_{\partial^* E} \varphi(x) \nu_E(x) \, d\mathcal{H}^{n-1}(x) \quad \text{for all } \varphi \in C_c^1(\mathbb{R}^n).$$

In particular, $P(E; A) = \mathcal{H}^{n-1}(A \cap \partial^* E)$ for every Borel set $A \subset \mathbb{R}^n$. We say that $x \in \mathbb{R}^n$ is a jump point of E if there exists $\nu \in S^{n-1}$ such that

$$E_{x,r} \xrightarrow{\text{loc}} H_{0,\nu}^+ \quad \text{as } r \to 0^+, \tag{2-13}$$

and we denote by $\partial^J E$ the set of *jump points* of E. Note that we always have $\partial^J E \subset E^{(1/2)} \subset \partial^e E$. In fact, if E is a set of locally finite perimeter and $x \in \partial^* E$, then (2-13) holds with $\nu = -\nu_E(x)$, so that $\partial^* E \subset \partial^J E$. Summarizing, if E is a set of locally finite perimeter, we have

$$\partial^* E \subset \partial^J E \subset E^{(1/2)} \subset \partial^e E \tag{2-14}$$

and, moreover, by *Federer's theorem* [Ambrosio et al. 2000, Theorem 3.61; Maggi 2012, Theorem 16.2],

$$\mathcal{H}^{n-1}(\partial^e E \setminus \partial^* E) = 0,$$

so that $\partial^e E$ is locally $\mathcal{H}^{n-1}$-rectifiable in $\mathbb{R}^n$. We shall need on several occasions to use the following very fine criterion for finite perimeter, known as *Federer's criterion* [1969, 4.5.11] (see also [Evans and Gariepy 1992, Section 5.11, Theorem 1]): if E is a Lebesgue measurable set in $\mathbb{R}^n$ such that $\partial^e E$ is locally $\mathcal{H}^{n-1}$-finite, then E is a set of locally finite perimeter.

Given a Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ and an open set $\Omega \subset \mathbb{R}^n$ we define the *total variation of f in Ω* as

$$|Df|(\Omega) = \sup \left\{ \int_\Omega f(x) \operatorname{Div} T(x) \, dx : T \in C_c^1(\Omega; \mathbb{R}^n), \ |T| \leq 1 \right\}.$$

We say that $f \in BV(\Omega)$ if $|Df|(\Omega) < \infty$ and $f \in L^1(\Omega)$, and that $f \in BV_{\text{loc}}(\Omega)$ if $f \in BV(\Omega')$ for every open set Ω' compactly contained in Ω. If $f \in BV_{\text{loc}}(\mathbb{R}^n)$ then the distributional derivative Df of f is an $\mathbb{R}^n$-valued Radon measure. Note in particular that E is a set of locally finite perimeter if and only if $1_E \in BV_{\text{loc}}(\mathbb{R}^n)$, and that in this case $\mu_E = -D1_E$. Sets of finite perimeter and functions of bounded variation are related by the fact that, if $f \in BV_{\text{loc}}(\mathbb{R}^n)$, then, for a.e. $t \in \mathbb{R}$, $\{f > t\}$ is a set of finite perimeter, and the *coarea formula*,

$$\int_{\mathbb{R}} P(\{f > t\}; G) \, dt = |Df|(G), \tag{2-15}$$

holds (as an identity in $[0, \infty]$) for every Borel set $G \subset \mathbb{R}^n$. If $f \in BV_{\text{loc}}(\mathbb{R}^n)$, then the Radon–Nikodym decomposition of Df with respect to $\mathcal{H}^n$ is denoted by $Df = D^a f + D^s f$, where $D^s f$ and $\mathcal{H}^n$ are mutually singular, and where $D^a f \ll \mathcal{H}^n$. The density of $D^a f$ with respect to $\mathcal{H}^n$ is by convention denoted as ∇f, so that $\nabla f \in L^1(\Omega; \mathbb{R}^n)$ with $D^a f = \nabla f \, d\mathcal{H}^n$. Moreover, for a.e. $x \in \mathbb{R}^n$, $\nabla f(x)$ is the approximate differential of f at x. If $f \in BV_{\text{loc}}(\mathbb{R}^n)$, then S_f is countably $\mathcal{H}^{n-1}$-rectifiable with $\mathcal{H}^{n-1}(S_f \setminus J_f) = 0$,

$[f] \in L^1_{\text{loc}}(\mathcal{H}^{n-1} \llcorner J_f)$, and the $\mathbb{R}^n$-valued Radon measure $D^j f$, defined as

$$D^j f = [f] \nu_f \, d\mathcal{H}^{n-1} \llcorner J_f,$$

is called the *jump part of Df*. Since $D^a f$ and $D^j f$ are mutually singular, by setting $D^c f = D^s f - D^j f$ we come to the canonical decomposition of Df into the sum $D^a f + D^j f + D^c f$. The $\mathbb{R}^n$-valued Radon measure $D^c f$ is called the *Cantorian part* of Df. It has the distinctive property that $|D^c f|(M) = 0$ if M is σ-finite with respect to $\mathcal{H}^{n-1}$. We shall often need to use (in combination with (2-10) and (2-12)) the following localization property of Cantorian derivatives.

Lemma 2.2. *If $v \in BV(\mathbb{R}^n)$, then $|D^c v|(\{v^\wedge = 0\}) = 0$. In particular, if $f, g \in BV(\mathbb{R}^n)$ and $f = g$ $\mathcal{H}^n$-a.e. on a Borel set E, then $D^c f \llcorner E^{(1)} = D^c g \llcorner E^{(1)}$.*

Proof. Step one: Let $v \in BV(\mathbb{R}^n)$, and let $K \subset S_v^c$ be a concentration set for $D^c v$ that is $\mathcal{H}^n$-negligible. By the coarea formula,

$$|D^c v|(\{v^\wedge = 0\}) = |D^c v|(K \cap \{v^\wedge = 0\}) = |Dv|(K \cap \{v^\wedge = 0\})$$

$$= \int_{\mathbb{R}} \mathcal{H}^{n-2}(K \cap \{v^\wedge = 0\} \cap \partial^*\{v > t\}) \, dt$$

$$= \int_{\mathbb{R}} \mathcal{H}^{n-2}(K \cap \{\tilde{v} = 0\} \cap \partial^*\{v > t\}) \, dt = 0 \qquad \text{(by } v^\wedge = v^\vee \text{ on } S_v^c\text{),}$$

where in the last identity we have noticed that $\{\tilde{v} = 0\} \cap \partial^*\{v > t\} \cap S_v^c = \varnothing$ if $t \neq 0$.

Step two: Let $f, g \in BV(\mathbb{R}^n)$ with $f = g$ $\mathcal{H}^n$-a.e. on a Borel set E. Let $v = f - g$ so that $v \in BV(\mathbb{R}^n)$. Since $v = 0$ on E, we easily see that $E^{(1)} \subset \{\tilde{v} = 0\}$. Thus $|D^c v|(E^{(1)}) = 0$, by step one. $\qquad \square$

Lemma 2.3. *If $f, g \in BV(\mathbb{R}^n)$, E is a set of finite perimeter, and $f = 1_E g$, then*

$$\nabla f = 1_E \nabla g \quad \mathcal{H}^n\text{-a.e. on } \mathbb{R}^n, \tag{2-16}$$

$$D^c f = D^c g \llcorner E^{(1)}, \tag{2-17}$$

$$S_f \cap E^{(1)} = S_g \cap E^{(1)}. \tag{2-18}$$

Proof. Since $f = g$ on E, by (2-12) we find that $\nabla f = \nabla g$ $\mathcal{H}^n$-a.e. on E; since $f = 0$ on $\mathbb{R}^n \setminus E$, again by (2-12) we find that $\nabla f = 0$ $\mathcal{H}^n$-a.e. on $\mathbb{R}^n \setminus E$; this proves (2-16). For the same reasons, but this time exploiting Lemma 2.2 in place of (2-12), we see that $D^c f \llcorner E^{(1)} = D^c g \llcorner E^{(1)}$ and that $D^c f \llcorner (\mathbb{R}^n \setminus E)^{(1)} = D^c f \llcorner E^{(0)} = 0$; since $\partial^e E$ is locally $\mathcal{H}^{n-2}$-rectifiable, and thus $|D^c f|$-negligible, we come to (2-17). Finally, (2-18) is an immediate consequence of (2-10). $\qquad \square$

Given a Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ we say that f is a function of *generalized bounded variation* on $\mathbb{R}^n$, $f \in GBV(\mathbb{R}^n)$, if $\psi \circ f \in BV_{\text{loc}}(\mathbb{R}^n)$ for every $\psi \in C^1(\mathbb{R})$ with $\psi' \in C^0_c(\mathbb{R})$, or, equivalently, if $\tau_M(f) \in BV_{\text{loc}}(\mathbb{R}^n)$ for every $M > 0$, where τ_M was defined in (2-5). Note that, if $f \in GBV(\mathbb{R}^n)$, then we do not require that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, so that the distributional derivative Df of f may even fail to be defined. Nevertheless, the structure theory of *BV*-functions holds for *GBV*-functions too. Indeed, if $f \in GBV(\mathbb{R}^n)$, then — see [Ambrosio et al. 2000, Theorem 4.34] — $\{f > t\}$ is a set of finite

perimeter for a.e. $t \in \mathbb{R}$, f is approximately differentiable $\mathcal{H}^n$-a.e. on $\mathbb{R}^n$, S_f is countably $\mathcal{H}^{n-1}$-rectifiable and $\mathcal{H}^{n-1}$-equivalent to J_f, and the coarea formula (2-15) takes the form

$$\int_{\mathbb{R}} P(\{f > t\}; G) \, dt = \int_G |\nabla f| \, d\mathcal{H}^n + \int_{G \cap S_f} [f] \, d\mathcal{H}^{n-1} + |D^c f|(G) \tag{2-19}$$

for every Borel set $G \subset \mathbb{R}^n$, where $|D^c f|$ denotes the Borel measure on $\mathbb{R}^n$ defined as the least upper bound of the Radon measures $|D^c(\tau_M(f))|$; and, in fact,

$$|D^c f|(G) = \lim_{M \to \infty} |D^c(\tau_M(f))|(G) = \sup_{M>0} |D^c(\tau_M(f))|(G) \tag{2-20}$$

whenever G is a Borel set in $\mathbb{R}^n$; see [Ambrosio et al. 2000, Definition 4.33].

3. Characterization of equality cases and barycenter functions

We now prove the results presented in Section 1D. In Section 3A, Theorem 3.1, we obtain a formula for the perimeter of a set whose sections are segments, which is then applied in Section 3B to study barycenter functions of such sets and prove Theorem 1.7. Sections 3C and 3D contain the proof of Theorem 1.9 concerning the characterization of equality cases in terms of barycenter functions, while Theorem 1.13 is proved in Section 3E.

3A. Sets with segments as sections. Given $u : \mathbb{R}^{n-1} \to \mathbb{R} \cup \{\pm\infty\}$, let $\Sigma_u = \{x \in \mathbb{R}^n : \boldsymbol{q}x > u(\boldsymbol{p}x)\}$ and $\Sigma^u = \{x \in \mathbb{R}^n : \boldsymbol{q}x < u(\boldsymbol{p}x)\}$, respectively, denote the epigraph and the subgraph of u. As proved in [Cagnetti et al. 2013, Proposition 3.1], Σ_u is a set of locally finite perimeter if and only if $\tau_M(u) \in BV_{\text{loc}}(\mathbb{R}^{n-1})$ for every $M > 0$. (Note that this does not mean that $u \in GBV(\mathbb{R}^{n-1})$, as here u takes values in $\mathbb{R} \cup \{\pm\infty\}$.) Moreover, it is well known that if $u \in BV_{\text{loc}}(\mathbb{R}^{n-1})$ then, for every Borel set $G \subset \mathbb{R}^{n-1}$, the identity

$$P(\Sigma_u; G \times \mathbb{R}) = \int_G \sqrt{1 + |\nabla u|^2} \, d\mathcal{H}^{n-1} + \int_{G \cap S_u} [u] \, d\mathcal{H}^{n-2} + |D^c u|(G) \tag{3-1}$$

holds in $[0, \infty]$; see [Giaquinta et al. 1998b, Chapter 4, Sections 1.5 and 2.4]. In the study of equality cases for Steiner's inequality, thanks to Theorem A, we are concerned with sets E of the form $E = \Sigma_{u_1} \cap \Sigma^{u_2}$ corresponding to Lebesgue measurable functions u_1 and u_2 such that $u_1 \leq u_2$ on $\mathbb{R}^{n-1}$. A characterization of those pairs of functions u_1, u_2 corresponding to sets E of finite perimeter and volume is presented in Proposition 3.2. In Theorem 3.1, we provide instead a formula that is analogous to (3-1) for the perimeter of E in terms of u_1 and u_2 in the case that $u_1, u_2 \in GBV(\mathbb{R}^{n-1})$.

Theorem 3.1. *If $u_1, u_2 \in GBV(\mathbb{R}^{n-1})$ with $u_1 \leq u_2$, and $E = \Sigma_{u_1} \cap \Sigma^{u_2}$ has finite volume, then E is a set of locally finite perimeter and, for every Borel set $G \subset \mathbb{R}^{n-1}$,*

$$P(E; G \times \mathbb{R}) = \int_{G \cap \{u_1 < u_2\}} \sqrt{1 + |\nabla u_1|^2} \, d\mathcal{H}^{n-1} + \int_{G \cap \{u_1 < u_2\}} \sqrt{1 + |\nabla u_2|^2} \, d\mathcal{H}^{n-1} + |D^c u_1|(G \cap \{\tilde{u}_1 < \tilde{u}_2\})$$

$$+ |D^c u_2|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}) + \int_{G \cap (S_{u_1} \cup S_{u_2})} \min\{2(\tilde{u}_2 - \tilde{u}_1), [u_1] + [u_2]\} \, d\mathcal{H}^{n-2}, \tag{3-2}$$

where this identity holds in $[0, \infty]$, and with the convention that $\tilde{u}_2 - \tilde{u}_1 = 0$ when $\tilde{u}_2 = \tilde{u}_1 = +\infty$.
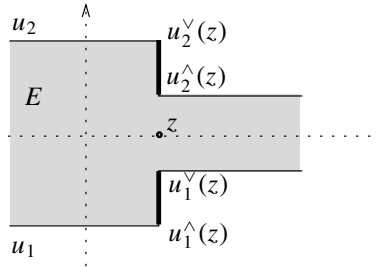
Figure 7. The inclusion (3-3).

If $E = \Sigma_{u_1} \cap \Sigma^{u_2}$ is of locally finite perimeter, then it is not necessarily true that $u_1, u_2 \in GBV(\mathbb{R}^{n-1})$. The regularity of u_1 and u_2 is, in fact, quite minimal, and completely degenerates as we approach the set where u_1 and u_2 coincide.

Proposition 3.2. *Let $u_1, u_2 : \mathbb{R}^{n-1} \to \mathbb{R}$ be Lebesgue measurable functions with $u_1 \leq u_2$ on $\mathbb{R}^{n-1}$. Then $E = \Sigma_{u_1} \cap \Sigma^{u_2}$ is of finite perimeter with $0 < |E| < \infty$ if and only if $v = u_2 - u_1 \in BV(\mathbb{R}^{n-1})$, $v \neq 0$, $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, $\{u_2 > t > u_1\}$ is of finite perimeter for a.e. $t \in \mathbb{R}$, and $f \in L^1(\mathbb{R})$ for $f(t) = P(\{u_2 > t > u_1\})$, $t \in \mathbb{R}$. In particular,*

$$\int_{\mathbb{R}} P(\{u_2 > t > u_1\}) \, dt \leq P(E),$$
$$|Dv|(\mathbb{R}^{n-1}) \leq P(F[v]),$$
$$\mathcal{H}^{n-1}(\{v > 0\}) \leq \tfrac{1}{2} P(F[v]).$$

Moreover (see Figure 7),

$$(\partial^e E)_z \subset [u_1^\wedge(z), u_1^\vee(z)] \cup [u_2^\wedge(z), u_2^\vee(z)] \quad \text{for all } z \in \mathbb{R}^{n-1}, \tag{3-3}$$

and

$$(S_{u_1} \cup S_{u_2}) \setminus (\{u_2^\vee = u_1^\vee\} \cap \{u_2^\wedge = u_1^\wedge\}) \tag{3-4}$$

is countably $\mathcal{H}^{n-2}$-rectifiable, with $\{v^\vee = 0\} \subseteq \{u_2^\vee = u_1^\vee\} \cap \{u_2^\wedge = u_1^\wedge\}$.

Proof. We first note that, if we set $E(t) = \{z \in \mathbb{R}^{n-1} : (z, t) \in E\}$, then we have $E(t) = \{u_1 < t < u_2\}$ for every $t \in \mathbb{R}$, and that, by Fubini's theorem, E has finite volume if and only if $v \in L^1(\mathbb{R}^{n-1})$; if these hold, then $|E| = \int_{\mathbb{R}^{n-1}} v$.

Step one: Let us assume that E has finite perimeter and that $0 < |E| < \infty$; in particular, $v \in L^1(\mathbb{R}^{n-1})$. By Steiner's inequality, $F[v]$ has finite perimeter. By [Maggi 2012, Proposition 19.22], since $|F[v] \cap \{x_n > 0\}|$ equals $\int_{\mathbb{R}^{n-1}} \frac{1}{2} v = \frac{1}{2}|E| > 0$, we have that

$$\tfrac{1}{2} P(F[v]) \geq P(F[v]; \{x_n > 0\}) \geq \mathcal{H}^{n-1}(F[v]^{(1)} \cap \{x_n = 0\}) = \mathcal{H}^{n-1}(\{v > 0\}).$$

If $T \in C_c^1(\mathbb{R}^{n-1}; \mathbb{R}^{n-1})$ with $\sup_{\mathbb{R}^{n-1}} |T| \leq 1$, and we set $S \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$ to be $S(x) = (T(\boldsymbol{p}x), 0)$, then by Fubini's theorem and Steiner's inequality we find that

$$\int_{\mathbb{R}^{n-1}} v(z) \operatorname{Div} T(z) \, dz = \int_{F[v]} \operatorname{Div} S \leq P(F[v]) \leq P(E).$$

Hence, $v \in BV(\mathbb{R}^{n-1})$ with $|Dv|(\mathbb{R}^{n-1}) \leq P(F[v])$. If $w_h \in C_c^1(\mathbb{R}^n)$ with $w_h \to 1_E$ in $L^1(\mathbb{R}^n)$ and $|Dw_h|(\mathbb{R}^n) \to P(E)$ as $h \to \infty$, then $w_h(\cdot, t) \to 1_{E(t)}$ in $L^1(\mathbb{R}^{n-1})$ for a.e. $t \in \mathbb{R}$, and, therefore,

$$\int_{E(t)} \mathrm{Div}\, T = \lim_{h \to \infty} \int_{\mathbb{R}^{n-1}} w_h \mathrm{Div}\, T = -\lim_{h \to \infty} \int_{\mathbb{R}^{n-1}} T \cdot \nabla w_h \leq \lim_{h \to \infty} \int_{\mathbb{R}^{n-1}} |\nabla w_h(z, t)|\, dz.$$

Hence, by Fatou's lemma,

$$\int_{\mathbb{R}} \sup \left\{ \left| \int_{E(t)} \mathrm{Div}\, T \right| : T \in C_c^1(\mathbb{R}^{n-1}; \mathbb{R}^{n-1}),\, \sup_{\mathbb{R}^{n-1}} |T| \leq 1 \right\} dt \leq \liminf_{h \to \infty} \int_{\mathbb{R}^n} |\nabla w_h| = P(E),$$

so that $E(t)$ is of finite perimeter for a.e. $t \in \mathbb{R}$, and $\int_{\mathbb{R}} P(E(t))\, dt \leq P(E)$, as required.

Step two: We now show the converse implication. To this end let $\varphi \in C_c^1(\mathbb{R}^n)$, then

$$\int_E \partial_n \varphi = \int_{\mathbb{R}^{n-1}} \varphi(z, u_2(z)) - \varphi(z, u_1(z))\, dz \leq 2 \sup_{\mathbb{R}^n} |\varphi| \mathcal{H}^{n-1}(\{v > 0\}),$$

while

$$\int_E \nabla_z \varphi = \int_{\mathbb{R}} dt \int_{E(t)} \nabla_z \varphi(z, t)\, dz = \int_{\mathbb{R}} dt \int_{\partial^* E(t)} \varphi(z, t) \nu_{E(t)}(z)\, d\mathcal{H}^{n-2}(z) \leq \sup_{\mathbb{R}^n} |\varphi| \int_{q(\mathrm{spt}\,\varphi)} P(E(t))\, dt.$$

If we set $f(t) = P(E(t))$, then we have just proved

$$\left| \int_E \nabla \varphi \right| \leq \sup_{\mathbb{R}^n} |\varphi| \left(2\mathcal{H}^{n-1}(\{v > 0\}) + \|f\|_{L^1(\mathbb{R})} \right),$$

so that E has finite perimeter.

Step three: For every $x \in \mathbb{R}^n$ and $r > 0$ we have

$$\mathcal{H}^n(E \cap C_{x,r}) = \int_{qx-r}^{qx+r} \mathcal{H}^{n-1}(D_{px,r} \cap \{u_1 < s\} \cap \{u_2 > s\})\, ds.$$

If $qx > u_2^\vee(px)$, then given $t \in (u_2^\vee(px), qx)$ and $r < qx - t$ we find that

$$\mathcal{H}^n(E \cap C_{x,r}) \leq 2r\mathcal{H}^{n-1}(D_{px,r} \cap \{u_2 > t\}) = o(r^n),$$

so that $x \in E^{(0)}$. By a similar argument, we show that

$$\{x \in \mathbb{R}^n : qx > u_2^\vee(px)\} \cup \{x \in \mathbb{R}^n : qx < u_1^\wedge(px)\} \subset E^{(0)},$$
$$\{x \in \mathbb{R}^n : u_1^\vee(px) < qx < u_2^\wedge(px)\} \subset E^{(1)}.$$

We thus conclude that, if $x \in \partial^e E$, then $u_1^\wedge(px) \leq qx \leq u_2^\vee(px)$ and either $qx \leq u_1^\vee(px)$ or $qx \geq u_2^\wedge(px)$.

Step four: Let I be a countable dense subset of $\mathbb{R}$ such that $\{u_1 < t < u_2\}$ is of finite perimeter for every $t \in I$. We claim that

$$\{u_2^\wedge > u_1^\wedge\} \cap S_{u_1} \subset \bigcup_{t \in I} \partial^e \{u_2 > t > u_1\}. \tag{3-5}$$

Indeed, if $\min\{u_2^\wedge(z), u_1^\vee(z)\} > t > u_1^\wedge(z)$, then

$$\theta(\{u_2 > t\}, z) = 1, \quad \theta^*(\{u_1 < t\}, z) > 0, \quad \theta_*(\{u_1 < t\}, z) < 1,$$

which implies that $\theta^*(\{u_1 < t < u_2\}, z) > 0$ and $\theta_*(\{u_1 < t < u_2\}, z) < 1$, and thus (3-5). In particular, $\{u_2^\wedge > u_1^\wedge\} \cap S_{u_1}$ is countably $\mathcal{H}^{n-2}$-rectifiable. By entirely similar arguments, one may check that the sets $\{u_2^\vee > u_1^\vee\} \cap S_{u_2}$, $S_{u_1}^c \cap S_{u_2}$, and $S_{u_1} \cap S_{u_2}^c$ are included in the set on the right-hand side of (3-5), and thus complete the proof of (3-4).

Step five: We prove that $\{v^\vee = 0\} \subseteq \{u_2^\vee = u_1^\vee\} \cap \{u_2^\wedge = u_1^\wedge\}$. Indeed from the general fact that $(f + g)^\vee \le f^\vee + g^\vee$, we obtain that $0 \le u_2^\vee - u_1^\vee \le (u_2 - u_1)^\vee = v^\vee$. At the same time,

$$0 \le u_2^\wedge - u_1^\wedge = (-u_1)^\vee - (-u_2)^\vee \le (-u_1 + u_2)^\vee = v^\vee. \qquad \square$$

Proof of Theorem 3.1. Step one: We first consider the case that $u_1, u_2 \in BV_{\text{loc}}(\mathbb{R}^{n-1})$. By [Giaquinta et al. 1998a, Section 4.1.5], Σ_{u_1} and Σ^{u_2} are of locally finite perimeter, with

$$\partial^* \Sigma_{u_1} \cap (S_{u_1}^c \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : \tilde{u}_1(\boldsymbol{p}x) = \boldsymbol{q}x\}, \tag{3-6}$$

$$\partial^* \Sigma_{u_1} \cap (S_{u_1} \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : u_1^\wedge(\boldsymbol{p}x) < \boldsymbol{q}x < u_1^\vee(\boldsymbol{p}x)\}, \tag{3-7}$$

and, by similar arguments, with

$$\Sigma_{u_1}^{(1)} \cap (S_{u_1}^c \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : \tilde{u}_1(\boldsymbol{p}x) < \boldsymbol{q}x\}, \tag{3-8}$$

$$\Sigma_{u_1}^{(1)} \cap (S_{u_1} \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : u_1^\vee(\boldsymbol{p}x) < \boldsymbol{q}x\}, \tag{3-9}$$

$$(\Sigma^{u_2})^{(1)} \cap (S_{u_2}^c \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : \tilde{u}_2(\boldsymbol{p}x) > \boldsymbol{q}x\}, \tag{3-10}$$

$$(\Sigma^{u_2})^{(1)} \cap (S_{u_2} \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{x \in \mathbb{R}^n : u_2^\wedge(\boldsymbol{p}x) > \boldsymbol{q}x\}. \tag{3-11}$$

Let us now recall that, by [Maggi 2012, Theorem 16.3], if F_1, F_2 are sets of locally finite perimeter, then

$$\partial^*(F_1 \cap F_2) =_{\mathcal{H}^{n-1}} (F_1^{(1)} \cap \partial^* F_2) \cup (F_2^{(1)} \cap \partial^* F_1) \cup (\partial^* F_1 \cap \partial^* F_2 \cap \{\nu_{F_1} = \nu_{F_2}\}); \tag{3-12}$$

moreover, if $F_1 \subset F_2$, then $\nu_{F_1} = \nu_{F_2}$ $\mathcal{H}^{n-1}$-a.e. on $\partial^* F_1 \cap \partial^* F_2$. Since $u_1 \le u_2$ implies $\Sigma_{u_2} \subset \Sigma_{u_1}$, and $\Sigma^{u_2} = \mathbb{R}^n \setminus \Sigma_{u_2}$, so that $\mu_{\Sigma_{u_2}} = -\mu_{\Sigma^{u_2}}$, we thus find

$$\nu_{\Sigma_{u_1}} = -\nu_{\Sigma^{u_2}} \quad \mathcal{H}^{n-1}\text{-a.e. on } \partial^* \Sigma_{u_1} \cap \partial^* \Sigma^{u_2}. \tag{3-13}$$

By (3-12) and (3-13), since $E = \Sigma_{u_1} \cap \Sigma^{u_2}$ we find

$$\partial^* E =_{\mathcal{H}^{n-1}} \left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)} \right) \cup \left(\partial^* \Sigma^{u_2} \cap (\Sigma_{u_1})^{(1)} \right).$$

We now apply (3-6) to u_1 and (3-10) to u_2 to find

$$\left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)} \right) \cap ((S_{u_1}^c \cap S_{u_2}^c) \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{(z, \tilde{u}_1(z)) : z \in (S_{u_1}^c \cap S_{u_2}^c), \tilde{u}_1(z) < \tilde{u}_2(z)\}. \tag{3-14}$$

We combine (3-7) applied to u_1 and (3-10) applied to u_2 to find

$$\left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)} \right) \cap ((S_{u_1} \cap S_{u_2}^c) \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \left\{ (z, t) : z \in S_{u_1} \cap S_{u_2}^c, \, u_1^\wedge(z) < t < \min\{u_1^\vee(z), \tilde{u}_2(z)\} \right\}. \tag{3-15}$$

We combine (3-7) applied to u_1 and (3-11) applied to u_2 to find

$$\left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)} \right) \cap ((S_{u_1} \cap S_{u_2}) \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \left\{ (z, t) : z \in S_{u_1} \cap S_{u_2}, \, u_1^\wedge(z) < t < \min\{u_1^\vee(z), u_2^\wedge(z)\} \right\}. \tag{3-16}$$

We finally apply (3-6) to u_1 and (3-11) to u_2 to find

$$\left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)}\right) \cap ((S_{u_1}^c \cap S_{u_2}) \times \mathbb{R}) =_{\mathcal{H}^{n-1}} \{(z, \tilde{u}_1(z)) : z \in S_{u_1}^c \cap S_{u_2}, \ \tilde{u}_1(z) < u_2^{\wedge}(z)\}. \qquad (3\text{-}17)$$

This gives, by (3-1), and using (3-14) for the first two terms and (3-15) and (3-16) for the third term on the right-hand side,

$$\mathcal{H}^{n-1}\left(\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)} \cap (G \times \mathbb{R})\right)$$
$$= \int_{G \cap \{u_1 < u_2\}} \sqrt{1 + |\nabla u_1|^2} \, d\mathcal{H}^{n-1} + |D^c u_1|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}) + \int_{G \cap S_{u_1}} (\min\{u_1^{\vee}, u_2^{\wedge}\} - u_1^{\wedge})_+ \, d\mathcal{H}^{n-2},$$

where we have also used that, as a consequence of (3-17), we simply have

$$\mathcal{H}^{n-1}\left((\partial^* \Sigma_{u_1} \cap (\Sigma^{u_2})^{(1)}) \cap ((S_{u_1}^c \cap S_{u_2}) \times \mathbb{R})\right) = 0,$$

by [Federer 1969, 3.2.23]. Also, by exchanging the role of u_1 and u_2,

$$\mathcal{H}^{n-1}(\partial^* \Sigma^{u_2} \cap (\Sigma_{u_1})^{(1)} \cap (G \times \mathbb{R}))$$
$$= \int_{G \cap \{u_1 < u_2\}} \sqrt{1 + |\nabla u_2|^2} \, d\mathcal{H}^{n-1} + |D^c u_2|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}) + \int_{G \cap S_{u_2}} (u_2^{\vee} - \max\{u_2^{\wedge}, u_1^{\vee}\})_+ \, d\mathcal{H}^{n-2}.$$

In conclusion, we have proved

$$P(E; G \times \mathbb{R})$$
$$= \int_{G \cap \{u_1 < u_2\}} (\sqrt{1 + |\nabla u_1|^2} + \sqrt{1 + |\nabla u_2|^2}) \, d\mathcal{H}^{n-1} + |D^c u_1|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}) + |D^c u_2|(G \cap \{\tilde{u}_1 < \tilde{u}_2\})$$
$$+ \int_{G \cap (S_{u_1} \cup S_{u_2})} \left(\min\{u_1^{\vee}, u_2^{\wedge}\} - u_1^{\wedge}\right)_+ + \left(u_2^{\vee} - \max\{u_2^{\wedge}, u_1^{\vee}\}\right)_+ \, d\mathcal{H}^{n-2}. \qquad (3\text{-}18)$$

We thus deduce (3-2) by means of (3-18) and the identity

$$\min\{2(\tilde{u}_2 - \tilde{u}_1), [u_1] + [u_2]\} = \min\{u_2^{\vee} + u_2^{\wedge} - (u_1^{\vee} + u_1^{\wedge}), u_1^{\vee} - u_1^{\wedge} + u_2^{\vee} - u_2^{\wedge}\}$$
$$= u_2^{\vee} - u_1^{\wedge} + \min\{u_2^{\wedge} - u_1^{\vee}, u_1^{\wedge} - u_2^{\vee}\}$$
$$= u_2^{\vee} - u_1^{\wedge} + \min\{u_2^{\wedge}, u_1^{\vee}\} - \max\{u_2^{\wedge}, u_1^{\vee}\}$$
$$= (\min\{u_1^{\vee}, u_2^{\wedge}\} - u_1^{\wedge})_+ + (u_2^{\vee} - \max\{u_2^{\wedge}, u_1^{\vee}\})_+ \, .$$

This completes the proof of the theorem in the case that $u_1, u_2 \in BV_{\text{loc}}(\mathbb{R}^{n-1})$.

Step two: We now address the general case. If $u_1, u_2 \in GBV(\mathbb{R}^{n-1})$, then Σ_{u_1} and Σ^{u_2} are sets of locally finite perimeter, by [Cagnetti et al. 2013, Proposition 3.1], and thus E is of locally finite perimeter. We now prove (3-2). To this end, since (3-2) is an identity between Borel measures on $\mathbb{R}^{n-1}$, it suffices to consider the case that G is *bounded*. Given $M > 0$, let $E_M = \Sigma_{\tau_M(u_1)} \cap \Sigma^{\tau_M(u_2)}$. Since $\tau_M u_i \in BV_{\text{loc}}(\mathbb{R}^{n-1})$ for every $M > 0$, $i = 1, 2$, by step one we find that E_M is a set of locally finite perimeter, and that (3-2) holds on E_M with $\tau_M(u_1)$ and $\tau_M(u_2)$ in place of u_1 and u_2. We are thus going to complete the proof of

the theorem by showing that

$$P(E; G \times \mathbb{R}) = \lim_{M \to \infty} P(E_M; G \times \mathbb{R}), \tag{3-19}$$

$$\int_{G \cap \{u_1 < u_2\}} \sqrt{1 + |\nabla u_i|^2} \, d\mathcal{H}^{n-1} = \lim_{M \to \infty} \int_{G \cap \{\tau_M(u_1) < \tau_M(u_2)\}} \sqrt{1 + |\nabla \tau_M(u_i)|^2} \, d\mathcal{H}^{n-1}, \tag{3-20}$$

$$|D^c u_i|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}) = \lim_{M \to \infty} |D^c \tau_M(u_i)|(G \cap \{\widetilde{\tau_M(u_1)} < \widetilde{\tau_M(u_2)}\}), \tag{3-21}$$

and that

$$\int_{G \cap (S_{u_1} \cup S_{u_2})} \min\{2(\tilde{u}_2 - \tilde{u}_1), [u_1] + [u_2]\} \, d\mathcal{H}^{n-2}$$
$$= \lim_{M \to \infty} \int_{G \cap (S_{\tau_M(u_1)} \cup S_{\tau_M(u_2)})} \min\{2(\widetilde{\tau_M(u_2)} - \widetilde{\tau_M(u_1)}), [\tau_M(u_1)] + [\tau_M(u_2)]\} \, d\mathcal{H}^{n-2}. \tag{3-22}$$

Let us set $f_M(a, b) = \tau_M(b) - \tau_M(a)$ for $a, b \in \mathbb{R} \cup \{\pm\infty\}$. By (2-6), we can write the right-hand side of (3-22) as $\int_G h_M \, d\mathcal{H}^{n-2}$, where

$$h_M = 1_{S_{\tau_M(u_1)} \cup S_{\tau_M(u_2)}} \gamma(f_M(u_1^\vee, u_2^\vee), f_M(u_1^\wedge, u_2^\wedge), f_M(u_1^\wedge, u_1^\vee), f_M(u_2^\wedge, u_2^\vee))$$

for a function $\gamma : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ that is increasing in each of its arguments. Since, for every $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a \le b$, the quantity $f_M(a, b)$ is increasing in M, with

$$\lim_{M \to \infty} f_M(a, b) = \begin{cases} 0 & \text{if } a = b = +\infty \text{ or } a = b = -\infty, \\ b - a & \text{otherwise,} \end{cases}$$

we see that $\{S_{\tau_M(u_i)}\}_{M>0}$ is a monotone increasing family of sets whose union is S_{u_i}, $\{h_M\}_{M>0}$ is an increasing family of functions on $\mathbb{R}^{n-1}$, and that

$$\lim_{M \to \infty} h_M = 1_{S_{u_1} \cup S_{u_2}} \min\{2(\tilde{u}_2 - \tilde{u}_1), [u_1] + [u_2]\},$$

where the convention that $\tilde{u}_2 - \tilde{u}_1 = 0$ if $\tilde{u}_2 = \tilde{u}_1 = +\infty$ was also used; we have thus completed the proof of (3-22). Similarly, since

$$\{\widetilde{\tau_M(u_1)} < \widetilde{\tau_M(u_2)}\} = \{f_M(u_1^\vee, u_2^\vee) + f_M(u_1^\wedge, u_2^\wedge) > 0\} = \{f_M(u_1^\vee, u_2^\vee) > 0\} \cup \{f_M(u_1^\wedge, u_2^\wedge) > 0\},$$

$\{\{\widetilde{\tau_M(u_1)} < \widetilde{\tau_M(u_2)}\}\}_{M>0}$ is a monotone increasing family of sets whose union is $\{u_2^\vee > u_1^\vee\} \cup \{u_2^\wedge > u_1^\wedge\}$. Therefore, by definition of $|D^c u_i|$, we find, for $i = 1, 2$,

$$\lim_{M \to \infty} |D^c \tau_M u_i|(G \cap \{\widetilde{\tau_M(u_1)} < \widetilde{\tau_M(u_2)}\}) = |D^c u_i|(G \cap (\{u_2^\vee > u_1^\vee\} \cup \{u_2^\wedge > u_1^\wedge\}))$$
$$= |D^c u_i|(G \cap \{\tilde{u}_1 < \tilde{u}_2\}),$$

where in the last identity we used that $S_{u_1} \cup S_{u_2}$ is countably $\mathcal{H}^{n-2}$-rectifiable, and thus $|D^c u_i|$-negligible for $i = 1, 2$. This proves (3-21). Next, we note that

$$|\nabla \tau_M(u_i)| = 1_{\{|u_i| < M\}} |\nabla u_i| \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1},$$

so that (3-20) follows again by monotone convergence. By (3-2) applied to E_M, this shows in particular that the limit as $M \to \infty$ of $P(E_M; G \times \mathbb{R})$ exists in $[0, \infty]$. Thus, in order to prove (3-19) it suffices to show that $P(E; G \times \mathbb{R})$ is the limit of $P(E_{M_h}; G \times \mathbb{R})$ as $h \to \infty$, where $\{M_h\}_{h \in \mathbb{N}}$ has been chosen in such a way that

$$\lim_{h \to \infty} \mathcal{H}^{n-1}(E^{(1)} \cap \{|x_n| = M_h\}) = 0, \qquad \mathcal{H}^{n-1}(\partial^e E \cap \{|x_n| = M_h\}) = 0 \quad \text{for all } h \in \mathbb{N}. \tag{3-23}$$

(Notice that the choice of $\{M_h\}_{h \in \mathbb{N}}$ is possible because $|E| < \infty$ and $\mathcal{H}^{n-1} \llcorner \partial^e E$ is a Radon measure.) Indeed, by $E_M = E \cap \{|x_n| < M\}$, (3-23), and [Maggi 2012, Theorem 16.3], we have that

$$\partial^e E_{M_h} = \big(\{|x_n| < M_h\} \cap \partial^e E\big) \cup \big(\{|x_n| = M_h\} \cap E^{(1)}\big) \quad \text{for all } h \in \mathbb{N},$$

so that, by the first identity in (3-23), we find $P(E; G \times \mathbb{R}) = \lim_{h \to \infty} P(E_{M_h}; G \times \mathbb{R})$, as required. $\square$

In practice, we shall always apply Theorem 3.1 in situations where the sets under consideration are described in terms of their barycenter and slice length functions.

Corollary 3.3. *If $v \in (BV \cap L^\infty)(\mathbb{R}^{n-1}; [0, \infty))$, $b \in GBV(\mathbb{R}^{n-1})$, and*

$$W = W[v, b] = \{x \in \mathbb{R}^n : |\mathbf{q}x - b(\mathbf{p}x)| < \tfrac{1}{2} v(\mathbf{p}x)\}, \tag{3-24}$$

then $u_1 = b - \tfrac{1}{2} v \in GBV(\mathbb{R}^{n-1})$, $u_2 = b + \tfrac{1}{2} v \in GBV(\mathbb{R}^{n-1})$, W is a set of locally finite perimeter with finite volume, and for every Borel set $G \subset \mathbb{R}^{n-1}$ we have

$$P(W; G \times \mathbb{R}) = \int_{G \cap \{v > 0\}} \sqrt{1 + |\nabla(b + \tfrac{1}{2}v)|^2} + \sqrt{1 + |\nabla(b - \tfrac{1}{2}v)|^2} \, d\mathcal{H}^{n-1}$$

$$+ \int_{G \cap (S_v \cup S_b)} \min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\} \, d\mathcal{H}^{n-2}$$

$$+ |D^c(b + \tfrac{1}{2}v)|(G \cap \{\tilde{v} > 0\}) + |D^c(b - \tfrac{1}{2}v)|(G \cap \{\tilde{v} > 0\}), \tag{3-25}$$

where this identity holds in $[0, \infty]$.

Proof. It is easily seen that $(BV \cap L^\infty) + GBV \subset GBV$. By Theorem 3.1, $W = \Sigma_{u_1} \cap \Sigma^{u_2}$ is of locally finite perimeter, and $P(W; G \times \mathbb{R})$ can be computed by means of (3-2) for every Borel set $G \subset \mathbb{R}^{n-1}$. We are thus left to prove that, $\mathcal{H}^{n-2}$-a.e. on $S_{u_1} \cup S_{u_2}$,

$$\min\{2(\tilde{u}_2 - \tilde{u}_1), [u_1] + [u_2]\} = \min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\}. \tag{3-26}$$

On $J_{u_1} \cap J_{u_2} \cap \{\nu_{u_1} = \nu_{u_2}\}$, we have that

$$b^\vee = \tfrac{1}{2}(u_1^\vee + u_2^\vee), \quad v^\vee = \max\{u_2^\vee - u_1^\vee, u_2^\wedge - u_1^\wedge\},$$

$$b^\wedge = \tfrac{1}{2}(u_1^\wedge + u_2^\wedge), \quad v^\wedge = \min\{u_2^\vee - u_1^\vee, u_2^\wedge - u_1^\wedge\},$$

while on $J_{u_1} \cap J_{u_2} \cap \{\nu_{u_1} = -\nu_{u_2}\}$ we find

$$b^\vee = \max\{\tfrac{1}{2}(u_2^\vee + u_1^\wedge), \tfrac{1}{2}(u_2^\wedge + u_1^\vee)\}, \quad v^\vee = u_2^\vee - u_1^\wedge,$$

$$b^\wedge = \min\{\tfrac{1}{2}(u_2^\vee + u_1^\wedge), \tfrac{1}{2}(u_2^\wedge + u_1^\vee)\}, \quad v^\wedge = u_2^\wedge - u_1^\vee,$$

so that (3-26) is proved through an elementary case-by-case argument on $J_{u_1} \cap J_{u_2}$, and thus, $\mathcal{H}^{n-2}$-a.e. on $S_{u_1} \cap S_{u_2}$. At the same time, on $S_{u_1} \cap S_{u_2}^c$ we have

$$b^\vee = \tfrac{1}{2}(\tilde{u}_2 + u_1^\vee), \quad v^\vee = \tilde{u}_2 - u_1^\wedge,$$
$$b^\wedge = \tfrac{1}{2}(\tilde{u}_2 + u_1^\wedge), \quad v^\wedge = \tilde{u}_2 - u_1^\vee,$$

from which we easily deduce (3-26) on $S_{u_1} \cap S_{u_2}^c$; by symmetry, we see the validity of (3-26) on $S_{u_1}^c \cap S_{u_2}$, and thus conclude the proof of the corollary. $\qquad\square$

Corollary 3.4. *Let $v : \mathbb{R}^{n-1} \to [0, \infty)$ be Lebesgue measurable. Then, $F[v]$ is of finite perimeter and volume if and only if $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ and $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$. If these hold, let $F = F[v]$, then for every $z \in \mathbb{R}^{n-1}$ we have*

$$(-\tfrac{1}{2}v^\wedge(z), \tfrac{1}{2}v^\wedge(z)) \subset (F^{(1)})_z \subset [-\tfrac{1}{2}v^\wedge(z), \tfrac{1}{2}v^\wedge(z)], \tag{3-27}$$

$$\{t \in \mathbb{R} : \tfrac{1}{2}v^\wedge(z) < |t| < \tfrac{1}{2}v^\vee(z)\} \subset (\partial^e F)_z \subset \{t \in \mathbb{R} : \tfrac{1}{2}v^\wedge(z) \le |t| \le \tfrac{1}{2}v^\vee(z)\}, \tag{3-28}$$

while, for every Borel set $G \subset \mathbb{R}^{n-1}$,

$$P(F; G \times \mathbb{R}) = 2 \int_{G \cap \{v > 0\}} \sqrt{1 + |\tfrac{1}{2}\nabla v|^2} \, d\mathcal{H}^{n-1} + \int_{G \cap S_v} [v] \, d\mathcal{H}^{n-2} + |D^c v|(G). \tag{3-29}$$

Proof. By Proposition 3.2 and the coarea formula (2-15), we see that $F[v]$ is of finite perimeter if and only if $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ and $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$. By arguing as in step three of the proof of Proposition 3.2, we easily prove (3-27) and (3-28). Finally, by applying Theorem 3.1 to $u_2 = \tfrac{1}{2}v$ and $u_1 = -\tfrac{1}{2}v$, we prove (3-29) with $|D^c v|(G \cap \{\tilde{v} > 0\})$ in place of $|D^c v|(G)$. By Lemma 2.2, this concludes the proof of the corollary. $\qquad\square$

We close this section with the proof of Proposition 1.15.

Proof of Proposition 1.15. We want to prove that, if $\lambda \in [0, 1] \setminus \{\tfrac{1}{2}\}$ and

$$E = \left\{x \in \mathbb{R}^n : -\lambda v_2(\boldsymbol{p}x) - \tfrac{1}{2}v_1(\boldsymbol{p}x) \le \boldsymbol{q}x \le \tfrac{1}{2}v_1(\boldsymbol{p}x) + (1 - \lambda)v_2(\boldsymbol{p}x)\right\}, \tag{3-30}$$

then $E \in \mathcal{M}(v)$ and $\mathcal{H}^n\big(E \Delta (te_n + F[v])\big) > 0$ for every $t \in \mathbb{R}$. By Corollary 3.4,

$$P(F[v]) = 2 \int_{\mathbb{R}^{n-1}} \sqrt{1 + |\nabla(\tfrac{1}{2}v_1)|^2} + |D^s v_2|(\mathbb{R}^{n-1}). \tag{3-31}$$

At the same time, $E = W[v, b]$, where $b = (\tfrac{1}{2} - \lambda)v_2$. Since $D^s v_1 = 0$, $D^a v_2 = 0$, and

$$v^\vee + v^\wedge \ge [v] = [v_2] \ge 2[b] \quad \mathcal{H}^{n-2}\text{-a.e. on } \mathbb{R}^{n-1},$$

we easily find that

$$\nabla(b \pm \tfrac{1}{2}v) = \pm\nabla(\tfrac{1}{2}v_1) \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1},$$
$$\min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\} = [v_2] \quad \mathcal{H}^{n-2}\text{-a.e. on } \mathbb{R}^{n-1},$$
$$D^c(b + \tfrac{1}{2}v) = (1 - \lambda)D^c v_2,$$
$$D^c(b - \tfrac{1}{2}v) = -\lambda D^c v_2.$$

Since $S_b \cup S_v =_{\mathcal{H}^{n-2}} S_{v_2}$, we find $P(E) = P(F[v])$ by (3-31) and (3-25). At the same time,

$$\mathcal{H}^n\big(E \Delta (t e_n + F[v])\big) = 2 \int_{\{v > 0\}} |t - (\tfrac{1}{2} - \lambda) v_2| \, d\mathcal{H}^{n-1} \quad \text{for all } t \in \mathbb{R},$$

so that $\mathcal{H}^n\big(E \Delta (t e_n + F[v])\big) > 0$, as $\lambda \neq \tfrac{1}{2}$ and v_2 is nonconstant on $\{v > 0\}$. $\qquad\square$

3B. *A fine analysis of the barycenter function.* We now prove Theorem 1.7, which states in particular that $b_E 1_{\{v > \delta\}} \in GBV(\mathbb{R}^{n-1})$ whenever E is a v-distributed set of finite perimeter and $\{v > \delta\}$ is of finite perimeter. We first discuss some examples showing that this is the optimal degree of regularity we can expect for the barycenter. (Let us also recall that the regularity of barycenter functions in arbitrary codimension, but under "no vertical boundaries" and "no vanishing sections" assumptions, was addressed in [Barchiesi et al. 2013, Theorem 4.3].)

Remark 3.5. In the case $n = 2$, as will be clear from the proof of Theorem 1.7, conclusion (1-11) can be strengthened to $1_{\{v > \delta\}} b_E \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$. The localization on $\{v > \delta\}$ is necessary. Indeed, let us define $E \subset \mathbb{R}^2$ as

$$E = \bigcup_{h \in \mathbb{N}} \left\{ x \in \mathbb{R}^2 : \frac{1}{h+1} < \boldsymbol{p} x < \frac{1}{h}, \ |\boldsymbol{q} x - (-1)^h| < \frac{1}{h^2} \right\},$$

so that E has finite perimeter and volume, and has segments as sections. However,

$$b_E(z) = \sum_{h \in \mathbb{N}} (-1)^h 1_{((h+1)^{-1}, h^{-1})}(z), \quad z \in \mathbb{R},$$

so that $b_E \in L^\infty(\mathbb{R}) \setminus BV(\mathbb{R})$. We also note that, in the case $n \geq 3$, the use of generalized functions of bounded variation is necessary. For example, let $E_\alpha \subset \mathbb{R}^3$ be such that

$$E_\alpha = \bigcup_{h \in \mathbb{N}} \left\{ x \in \mathbb{R}^3 : \frac{1}{(h+1)^2} < |\boldsymbol{p} x| < \frac{1}{h^2}, \ |\boldsymbol{q} x - h^\alpha| < \frac{1}{2} \right\}, \quad \alpha > 0.$$

In this way, E_α always has finite perimeter and volume, with $v(z) = 1$ if $|z| < 1$ and

$$1_{\{v > \delta\}}(z) b_{E_\alpha}(z) = b_{E_\alpha}(z) = \sum_{h \in \mathbb{N}} 1_{((h+1)^{-2}, h^{-2})}(|z|) h^\alpha \quad \text{for all } z \in \mathbb{R}^2, \ 0 < \delta < 1.$$

In particular, $1_{\{v > \delta\}} b_{E_2} \in L^1(\mathbb{R}^2) \setminus BV(\mathbb{R}^2)$ and $1_{\{v > \delta\}} b_{E_4} \notin L^1_{\text{loc}}(\mathbb{R}^2)$. Hence, without truncation, $1_{\{v > \delta\}} b_E$ may either fail to be of bounded variation (even if it is locally summable), or it may just fail to be locally summable.

Before entering into the proof of Theorem 1.7, we shall need to prove that the momentum function m_E of a vertically bounded set E is of bounded variation; see Lemma 3.6 below. Given $E \subset \mathbb{R}^n$, we say that E is vertically bounded (by $M > 0$) if $E \subset_{\mathcal{H}^n} \{x \in \mathbb{R}^n : |\boldsymbol{q} x| < M\}$.

Lemma 3.6. *If $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ and E is a vertically bounded, v-distributed set of finite perimeter, then $m_E \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, where*

$$m_E(z) = \int_{E_z} t \, d\mathcal{H}^1(t) \quad \text{for all } z \in \mathbb{R}^{n-1}.$$

Proof. If E is vertically bounded by $M > 0$, then $v \in L^\infty(\mathbb{R}^{n-1})$, $|m_E| \le Mv$, and $m_E \in L^\infty(\mathbb{R}^{n-1})$. Moreover, $m_E \in BV(\mathbb{R}^{n-1})$ as, for every $\varphi \in C_c^1(\mathbb{R}^{n-1})$,

$$\int_{\mathbb{R}^{n-1}} m_E \nabla' \varphi \, d\mathcal{H}^{n-1} = \int_E \nabla'(\varphi(\boldsymbol{p}x)\boldsymbol{q}x) \, d\mathcal{H}^n(x) = \int_{\partial^* E} \varphi(\boldsymbol{p}x)\boldsymbol{q}x \, \boldsymbol{p}\nu_E(x) \, d\mathcal{H}^{n-1}(x) \le M \sup_{\mathbb{R}^{n-1}} |\varphi| P(E). \quad \square$$

Proof of Theorem 1.7. Step one: Let us decompose $z \in \mathbb{R}^{n-1}$ as $z = (z_1, z') \in \mathbb{R} \times \mathbb{R}^{n-2}$. For every fixed $z' \in \mathbb{R}^{n-2}$, $f : \mathbb{R}^{n-1} \to \mathbb{R}$, $G \subset \mathbb{R}^{n-1}$, and $E \subset \mathbb{R}^n$, we define

$$f^{z'} : \mathbb{R} \to \mathbb{R}, \quad f^{z'}(z_1) = f(z_1, z'),$$
$$G^{z'} = \{z_1 \in \mathbb{R} : (z_1, z) \in G\},$$
$$E^{z'} = \{(z_1, t) \in \mathbb{R}^2 : (z_1, z', t) \in E\}.$$

We now consider v and E as in the statement, and identify a set $I \subset (0, 1)$ such that $\mathcal{H}^1((0, 1) \setminus I) = 0$ and, if $\delta \in I$, then $\{v > \delta\}$ is a set of finite perimeter. We now fix $\delta \in I$, and consider a set $J \subset \mathbb{R}^{n-2}$ such that $\mathcal{H}^{n-2}(\mathbb{R}^{n-2} \setminus J) = 0$ and, for every $z' \in J$, $E^{z'}$ is a set of finite perimeter in $\mathbb{R}^2$ (hence, $v^{z'} \in BV(\mathbb{R})$) and $\{v > \delta\}^{z'} = \{v^{z'} > \delta\}$ is a set of finite perimeter in $\mathbb{R}$. Note that J depends on δ, and its existence is a consequence of Theorem C in Section 4D. As we shall see in step three, for every $z' \in J$,

$$\left| D(\tau_M(1_{\{v^{z'} > \delta\}} b_{E^{z'}})) \right|(\mathbb{R}) \le C(M, \delta)\{P(\{v^{z'} > \delta\}) + P(E^{z'})\}.$$

If we thus take into account that

$$(\tau_M(1_{\{v > \delta\}} b_E))^{z'} = \tau_M(1_{\{v^{z'} > \delta\}} b_{E^{z'}}),$$

we conclude that

$$\int_{\mathbb{R}^{n-2}} \left| D((\tau_M(1_{\{v > \delta\}} b_E))^{z'}) \right|(\mathbb{R}) \, d\mathcal{H}^{n-2}(z') \le C(M, \delta) \int_{\mathbb{R}^{n-2}} \{P(\{v^{z'} > \delta\}) + P(E^{z'})\} \, d\mathcal{H}^{n-2}(z')$$
$$\le C(M, \delta)\{P(\{v > \delta\}) + P(E)\},$$

where in the last step we have used [Maggi 2012, Proposition 14.5]. We can repeat this argument along each coordinate direction in $\mathbb{R}^{n-1}$ and combine it with [Ambrosio et al. 2000, Remark 3.104] to conclude that $\tau_M(1_{\{v > \delta\}} b_E) \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, with

$$\left| D(\tau_M(1_{\{v > \delta\}} b_E)) \right|(\mathbb{R}^{n-1}) \le C(M, \delta)\{P(\{v > \delta\}) + P(E)\}.$$

The proof of (1-11) will then be completed in the following two steps.

Step two: Let $n = 2$. We claim that $P(E^s) < \infty$ implies $v \in L^\infty(\mathbb{R})$, while $P(E) < \infty$ implies $b_E \in L^\infty(\{v > \sigma\})$ for every $\sigma > 0$. The first claim follows by Corollary 3.4: indeed, $P(E^s) < \infty$ implies $v \in BV(\mathbb{R})$ and thus, trivially, $v \in L^\infty(\mathbb{R})$. To prove the second claim, let us recall from step two in the proof of [Maggi 2012, Theorem 19.15] that if $a, b \in \mathbb{R}$ are such that $a \ne b$ and

$$\mathcal{H}^1(E_a^{(1)}) + \mathcal{H}^1(E_b^{(1)}) < \infty, \quad \mathcal{H}^1(E_a^{(1)} \cap E_b^{(1)}) = 0, \quad \mathcal{H}^1(\partial^* E_a^{(1)}) = \mathcal{H}^1(\partial^* E_b^{(1)}) = 0,$$

then one has

$$\mathcal{H}^1(E_a^{(1)}) + \mathcal{H}^1(E_b^{(1)}) \le P(E; \{a < x_1 < b\}). \tag{3-32}$$

Should b_E fail to be essentially bounded on $\{v > \sigma\}$ for some $\sigma > 0$, then we may construct a strictly increasing sequence $\{a_h\}_{h\in\mathbb{N}} \subset \mathbb{R}$ with $\sigma \leq \mathcal{H}^1(E_{a_h}^{(1)}) < \infty$, $\mathcal{H}^1(\partial^* E_{a_h}^{(1)}) = 0$, and $\mathcal{H}^1(E_{a_h}^{(1)} \cap E_{a_k}^{(1)}) = 0$ if $h \neq k$. Therefore, by (3-32), we would get

$$2\sigma \leq P(E; \{a_h < x_1 < a_{h+1}\}) \quad \text{for all } h \in \mathbb{N},$$

and thus conclude that $P(E) = +\infty$.

Step three: Let $v \in BV(\mathbb{R})$, let E be a v-distributed set of finite perimeter in $\mathbb{R}^2$ such that E_z is a segment for $\mathcal{H}^1$-a.e. $z \in \mathbb{R}$, and let $\delta > 0$ be such that $\{v > \delta\}$ is a set of finite perimeter in $\mathbb{R}$. According to step one, in order to complete the proof of (1-11) we are left to show that, if $M > 0$, then

$$\left| D(\tau_M(1_{\{v>\delta\}} b_E)) \right|(\mathbb{R}) \leq C(M, \delta)\{P(\{v > \delta\}) + P(E)\}. \tag{3-33}$$

By step two, $v \in L^\infty(\mathbb{R})$ and $b_E \in L^\infty(\{v > \delta\})$. In particular, E is vertically bounded above $\{v > \delta\}$, that is, there exists $L(\delta) > 0$ such that

$$E(\delta) = E \cap (\{v > \delta\} \times \mathbb{R}) \subset_{\mathcal{H}^2} \{x \in \mathbb{R}^2 : v(\boldsymbol{p}x) > \delta, |\boldsymbol{q}x| < L(\delta)\}. \tag{3-34}$$

Let us now set $v_\delta = 1_{\{v>\delta\}} v$. Since $\{v > \delta\}$ is of finite perimeter, we have

$$v_\delta \in (BV \cap L^\infty)(\mathbb{R}), \quad \{v_\delta > 0\} = \{v > \delta\}.$$

Concerning $E(\delta)$, we note that, since $\{v > \delta\} \times \mathbb{R}$ is of locally finite perimeter, then $E(\delta)$ is, at least, a v_δ-distributed set of locally finite perimeter such that $E(\delta)_z$ is a segment for $\mathcal{H}^1$-a.e. $z \in \mathbb{R}$. But, in fact, (3-34) implies $\{|x_n| > L(\delta)\} \subset E(\delta)^{(0)}$, while at the same time we have the inclusion

$$\partial^e E(\delta) \subset \left[\partial^e E \cap (\{v > \delta\}^{(1)} \times \mathbb{R})\right] \cup \left[(\partial^e\{v > \delta\} \times \mathbb{R}) \cap (E^{(1)} \cup \partial^e E)\right];$$

in particular, $E(\delta)$ is of finite perimeter by Federer's criterion, as

$$\mathcal{H}^{n-1}(\partial^e E(\delta)) \leq P(E; \{v > \delta\}^{(1)} \times \mathbb{R}) + 2L(\delta) P(\{v > \delta\}).$$

We now note that $b_{E(\delta)} = 1_{\{v>\delta\}} b_E \in L^\infty(\mathbb{R})$, with $P(E(\delta); \{v > \delta\}^{(1)} \times \mathbb{R}) \leq P(E)$; hence, (3-33) follows if we show that

$$\left| D(\tau_M(b_{E(\delta)})) \right|(\mathbb{R}) \leq C(M, \delta)\{P(\{v_\delta > 0\}) + P(E(\delta); \{v_\delta > 0\}^{(1)} \times \mathbb{R})\}$$

for every $M > 0$. It is now convenient to reset notation.

Step four: By step three, the proof of (1-11) will be completed by showing that, if $v \in (BV \cap L^\infty)(\mathbb{R})$ is such that, for some $\delta > 0$, $\{v > 0\} = \{v > \delta\}$ is a set of finite perimeter in $\mathbb{R}$, and E is a vertically bounded, v-distributed set of finite perimeter in $\mathbb{R}^2$ with $b_E \in L^\infty(\mathbb{R})$, then, for every $M > 0$,

$$\left| D(\tau_M(b_E)) \right|(\mathbb{R}) \leq C(M, \delta)\{P(\{v > 0\}) + P(E; \{v > 0\}^{(1)} \times \mathbb{R})\}. \tag{3-35}$$

We start by noting that, since E is vertically bounded, then by Lemma 3.6 we have $m_E \in (BV \cap L^\infty)(\mathbb{R})$. Moreover, if we set

$$w = \frac{1_{\{v>0\}}}{v} = \frac{1_{\{v>\delta\}}}{v},$$

then we have $w \in (BV \cap L^\infty)(\mathbb{R})$, and thus $b_E = w m_E \in (BV \cap L^\infty)(\mathbb{R})$. We now note that, since $\{v = 0\} \subset \{\tau_M(b_E) = 0\}$, we have $\{v = 0\}^{(1)} \subset \{\tau_M(b_E) = 0\}^{(1)}$; at the same time, a simple application of the coarea formula shows that

$$0 = \left| D(\tau_M(b_E)) \right| (\{\tau_M(b_E) = 0\}^{(1)}) \geq \left| D(\tau_M(b_E)) \right| (\{v = 0\}^{(1)}) = \left| D(\tau_M(b_E)) \right| (\{v > 0\}^{(0)}). \quad (3\text{-}36)$$

Moreover, since $\{v > 0\}$ is a set of finite perimeter, we know that $\partial^e\{v > 0\}$ is a finite set, so that

$$\left| D(\tau_M(b_E)) \right| (\partial^e\{v > 0\}) = \int_{S_{\tau_M(b_E)} \cap \partial^e\{v>0\}} [\tau_M(b_E)] \, d\mathcal{H}^0 \leq 2M P(\{v > 0\}), \quad (3\text{-}37)$$

where we have used that $[\tau_M(b_E)] \leq 2M$, since $|\tau_M(b_E)| \leq M$ on $\mathbb{R}^{n-1}$. By (3-36) and (3-37), in order to achieve (3-35) we are left to prove that

$$\left| D(\tau_M(b_E)) \right| (\{v > 0\}^{(1)}) \leq C(M, \delta) P(E; \{v > 0\}^{(1)} \times \mathbb{R}). \quad (3\text{-}38)$$

By (2-9) and since $\{v > \delta\} = \{v > 0\}$ we have

$$\{v^\wedge > 0\} \subset \{v > 0\}^{(1)} = \{v > \delta\}^{(1)} \subset \{v^\wedge \geq \delta\} \subset \{v^\wedge > 0\},$$

that is, $\{v > 0\}^{(1)} = \{v^\wedge > 0\}$. By applying Corollary 3.3 to $G = \{v > 0\}^{(1)} = \{v^\wedge > 0\}$,

$$P(E; \{v > 0\}^{(1)} \times \mathbb{R})$$
$$= \int_{\{v>0\}} \sqrt{1 + |(b_E + \tfrac{1}{2}v)'|^2} + \sqrt{1 + |(b_E - \tfrac{1}{2}v)'|^2} \, d\mathcal{H}^1 + \int_{\{v>0\}^{(1)} \cap (S_v \cup S_{b_E})} \min\{v^\vee + v^\wedge, \max\{[v], 2[b_E]\}\} \, d\mathcal{H}^0$$
$$+ |D^c(b_E + \tfrac{1}{2}v)|(\{v^\wedge > 0\} \cap \{\tilde{v} > 0\}) + |D^c(b_E - \tfrac{1}{2}v)|(\{v^\wedge > 0\} \cap \{\tilde{v} > 0\}). \quad (3\text{-}39)$$

Since $\{v^\wedge = 0\} = \{\tilde{v} = 0\} \cup \{v^\vee > 0 = v^\wedge\}$, where $\{v^\vee > 0 = v^\wedge\} \subset_{\mathcal{H}^0} J_v$, we find that $\{v^\wedge = 0\}$ is $|D^c f|$-equivalent to $\{\tilde{v} = 0\}$ for every $f \in BV_{\text{loc}}(\mathbb{R}^{n-1})$; hence,

$$|D^c(b_E \pm \tfrac{1}{2}v)|(\{v^\wedge > 0\} \cap \{\tilde{v} > 0\}) = |D^c(b_E \pm \tfrac{1}{2}v)|(\{v^\wedge > 0\}). \quad (3\text{-}40)$$

By (3-39), (3-40), the triangle inequality, and as $v^\wedge \geq \delta$ on $\{v > 0\}^{(1)} = \{v > \delta\}^{(1)}$,

$$P(E; \{v > 0\}^{(1)} \times \mathbb{R}) \geq 2 \int_{\{v>0\}} |b_E'| \, d\mathcal{H}^1 + 2 \int_{\{v>0\}^{(1)} \cap S_{b_E}} \min\{\delta, [b_E]\} \, d\mathcal{H}^0 + 2|D^c b_E|(\{v^\wedge > 0\}). \quad (3\text{-}41)$$

At the same time, by [Ambrosio et al. 2000, Theorem 3.99], for every $M > 0$ we have

$$|D(\tau_M(b_E))|(\{v > 0\}^{(1)}) = \int_{\{|b_E| < M\} \cap \{v>0\}} |b_E'| \, d\mathcal{H}^1 + |D^c b_E|(\{|\tilde{b}_E| < M\} \cap \{v > 0\}^{(1)})$$
$$+ \int_{S_{b_E} \cap \{b_E^\wedge < M\} \cap \{b_E^\vee > -M\} \cap \{v>0\}^{(1)}} \min\{M, b_E^\vee\} - \max\{-M, b_E^\wedge\} \, d\mathcal{H}^0. \quad (3\text{-}42)$$

As is easily seen by arguing on a case-by-case basis,

$$\min\{M, b_E^\vee\} - \max\{-M, b_E^\wedge\} \leq \max\left\{1, \frac{2M}{\delta}\right\} \min\{\delta, [b_E]\} \quad \text{on } S_{b_E}. \quad (3\text{-}43)$$

By combining (3-41), (3-42), and (3-43) we conclude the proof of (3-38), and thus of step four. The proof of (1-11) is now complete.

Step five: Since $\{v > \delta\}$ is of finite perimeter for a.e. $\delta > 0$, we find that $b_\delta = 1_{\{v>\delta\}} b_E \in GBV(\mathbb{R}^{n-1})$ for a.e. $\delta > 0$. In particular, b_δ is approximately differentiable at $\mathcal{H}^{n-1}$-a.e. $x \in \mathbb{R}^{n-1}$. Since $b_\delta = b_E$ on $\{v > \delta\}$, by (2-12) it follows that

$$\nabla b_E(x) = \nabla b_\delta(x) \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } x \in \{v > \delta\}. \tag{3-44}$$

By considering $\delta_h \to 0$ as $h \to \infty$ with $\{v > \delta_h\}$ of finite perimeter for every $h \in \mathbb{N}$, we find that b_E is approximately differentiable at $\mathcal{H}^{n-1}$-a.e. $x \in \{v > 0\}$. Since, trivially, b_E is approximately differentiable at *every* $x \in \{v = 0\}^{(1)}$ with $\nabla b_E(x) = 0$, we conclude that b_E is approximately differentiable at $\mathcal{H}^{n-1}$-a.e. $x \in \mathbb{R}^{n-1}$. By [Ambrosio et al. 2000, Theorem 4.34], for every Borel set $G \subset \mathbb{R}^{n-1}$ we have

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \partial^e\{b_\delta > t\}) \, dt = \int_G |\nabla b_\delta| \, d\mathcal{H}^{n-1} + \int_{G \cap S_{b_\delta}} [b_\delta] \, d\mathcal{H}^{n-2} + |D^c b_\delta|(G). \tag{3-45}$$

Let us note that, by (2-10), $[b_\delta] = [b_E]$ on $\{v > \delta\}^{(1)}$, and thus $S_{b_\delta} \cap \{v > \delta\}^{(1)} = S_{b_E} \cap \{v > \delta\}^{(1)}$. By (3-44) and by applying (3-45) to $G \cap \{v > \delta\}^{(1)}$, where $G \subset \mathbb{R}^{n-1}$ is a Borel set, we find

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \{v > \delta\}^{(1)} \cap \partial^e\{b_\delta > t\}) \, dt$$
$$= \int_{G \cap \{v>\delta\}} |\nabla b_E| \, d\mathcal{H}^{n-1} + \int_{G \cap S_{b_E} \cap \{v>\delta\}^{(1)}} [b_E] \, d\mathcal{H}^{n-2} + |D^c b_\delta|(G \cap \{v > \delta\}^{(1)}). \tag{3-46}$$

Since $\tau_M b_\delta = 1_{\{v>\delta\}} \tau_M b_\delta$, by applying Lemma 2.3 we find that, for every $G \subset \mathbb{R}^{n-1}$,

$$|D^c b_\delta|(G \cap \{v > \delta\}^{(1)}) = \lim_{M \to \infty} |D^c \tau_M b_\delta|(G \cap \{v > \delta\}^{(1)}) = \lim_{M \to \infty} |D^c \tau_M b_\delta|(G) = |D^c b_\delta|(G). \tag{3-47}$$

At the same time, since $\{v > \delta\} \cap \{b_\delta > t\} = \{v > \delta\} \cap \{b_E > t\}$ for every $t \in \mathbb{R}$, we have

$$\{v > \delta\}^{(1)} \cap \partial^e\{b_\delta > t\} = \{v > \delta\}^{(1)} \cap \partial^e\{b_E > t\} \quad \text{for all } t \in \mathbb{R},$$

and thus

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \{v > \delta\}^{(1)} \cap \partial^e\{b_\delta > t\}) \, dt = \int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \{v > \delta\}^{(1)} \cap \partial^e\{b_E > t\}) \, dt.$$

If we now set $\delta = \delta_h$ in (3-46) and then let $h \to \infty$, then since

$$\{v^\wedge > 0\} = \bigcup_{h \in \mathbb{N}} \{v > \delta_h\}^{(1)} \tag{3-48}$$

(which follows by (2-9)), by (3-47), and thanks to the definition (1-13) of $|D^c b_E|^+$, we find that (1-12) holds for every Borel set $G \subset \{v^\wedge > 0\}$, as required. We have thus completed the proof of Theorem 1.7. $\square$

3C. *Characterization of equality cases, part one.* In this section we prove the necessary conditions for equality cases in Steiner's inequality stated in Theorem 1.9. The proof requires the following simple lemma.

Lemma 3.7. *If μ and ν are $\mathbb{R}^{n-1}$-valued Radon measures on $\mathbb{R}^{n-1}$, then*

$$2|\mu|(G) \le |\nu + \mu|(G) + |\nu - \mu|(G) \tag{3-49}$$

for every Borel set $G \subset \mathbb{R}^{n-1}$. Moreover, equality holds in (3-49) *for every bounded Borel set $G \subset \mathbb{R}^{n-1}$ if and only if there exists a Borel function $f : \mathbb{R}^{n-1} \to [-1, 1]$ with*

$$\nu(G) = \int_G f \, d\mu \quad \text{for every bounded Borel set } G \subset \mathbb{R}^{n-1}.$$

Proof. The validity of (3-49) follows immediately from the fact that, if G is a Borel set in $\mathbb{R}^{n-1}$, then $|\mu|(G)$ is the supremum of the sums $\sum_{h \in \mathbb{N}} |\mu(G_h)|$ over partitions $\{G_h\}_{h \in \mathbb{N}}$ of G into bounded Borel sets. From the same fact, we immediately deduce that $|\nu + \mu|(G) = |\nu - \mu|(G) = |\nu|(G)$ whenever $|\mu|(G) = 0$; therefore, if G is such that $|\mu|(G) = 0$ and (3-49) holds as an equality, then $|\nu|(G) = 0$. In particular, if equality holds in (3-49) for every bounded Borel set $G \subset \mathbb{R}^{n-1}$, then $|\nu|$ is absolutely continuous with respect to $|\mu|$. By the Radon–Nikodym theorem we have that $\nu = g \, d|\mu|$ for a $|\mu|$-measurable function $g : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$, as well as $\mu = h \, d|\mu|$ for a $|\mu|$-measurable function $h : \mathbb{R}^{n-1} \to S^{n-2}$. In particular, $\nu \pm \mu = (g \pm h) \, d|\mu|$, and thus, since equality holds in (3-49),

$$2|\mu|(G) = |\nu + \mu|(G) + |\nu - \mu|(G) = \int_G |g + h| \, d|\mu| + \int_G |g - h| \, d|\mu|$$

for every Borel set $G \subset \mathbb{R}^{n-1}$, which gives

$$|g + h| + |h - g| = 2 = 2|h| \quad |\mu|\text{-a.e. on } \mathbb{R}^{n-1}.$$

Thus, there exists $\lambda : \mathbb{R}^{n-1} \to [0, \infty)$ such that $(h - g) = \lambda(g + h)$ $|\mu|$-a.e. on $\mathbb{R}^{n-1}$, i.e.,

$$g = \frac{1 - \lambda}{1 + \lambda} h \quad |\mu|\text{-a.e. on } \mathbb{R}^{n-1}.$$

This proves that $\nu = f \, d\mu$, where $f = (1 - \lambda)/(1 + \lambda)$. By Borel regularity of $|\mu|$, we can assume without loss of generality that f is Borel measurable. The proof is complete. $\square$

Proof of Theorem 1.9 (necessary conditions). Let $E \in \mathcal{M}(\nu)$. By Theorem A, we have that E_z is $\mathcal{H}^1$-equivalent to a segment for $\mathcal{H}^{n-1}$-a.e. $z \in \mathbb{R}^{n-1}$, which is (1-15). As a consequence, by Theorem 1.7, we have $b_\delta = 1_{\{v > \delta\}} b_E \in GBV(\mathbb{R}^{n-1})$ whenever $\{v > \delta\}$ is of finite perimeter. Let us set

$$I = \{\delta > 0 : \{v > \delta\} \text{ and } \{v < \delta\} \text{ are sets of finite perimeter}\}, \tag{3-50}$$

$$J_\delta = \{M > 0 : \{b_\delta < M\} \text{ and } \{b_\delta > -M\} \text{ are sets of finite perimeter}\}, \tag{3-51}$$

and note that $\mathcal{H}^1((0, \infty) \setminus I) = 0$ since $v \in BV(\mathbb{R}^{n-1})$, and that $\mathcal{H}^1((0, \infty) \setminus J_\delta) = 0$ for every $\delta \in I$, as $b_\delta \in GBV(\mathbb{R}^{n-1})$ whenever $\delta \in I$. By taking total variations in (1-18), we find $2|D^c(\tau_M b_\delta)|(G) \leq |D^c v|(G)$ for every bounded Borel set $G \subset \mathbb{R}^{n-1}$. By letting first $M \to \infty$ (in J_δ) and then $\delta \to 0$ (in I) we prove (1-19). Let us also note that (1-20) is an immediate corollary of (1-12) and (1-19), once (1-16) and (1-17) have been proved. Summarizing, these remarks show that we only need to prove the validity of (1-16), (1-17), and (1-18) (for $\delta \in I$ and $M \in J_\delta$) in order to complete the proof of the necessary conditions for equality cases. This is accomplished in various steps.

Step one: Let us fix $\delta, L \in I$ and $M \in J_\delta$, and set

$$\Sigma_{\delta, L, M} = \{\delta < v < L\} \cap \{|b_E| < M\} = \{|b_\delta| < M\} \cap \{\delta < v < L\},$$

so that $\Sigma_{\delta,L,M}$ is a set of finite perimeter. Since $\tau_M b_\delta \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ (see the end of step one in the proof of Theorem 1.7), $1_{\Sigma_{\delta,L,M}} \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, and $\tau_M b_\delta = b_\delta = b_E$ on $\Sigma_{\delta,L,M}$, we have

$$b_{\delta,L,M} = 1_{\Sigma_{\delta,L,M}} b_E \in (BV \cap L^\infty)(\mathbb{R}^{n-1}).$$

We now claim that there exists a Borel function $f_{\delta,L,M} : \mathbb{R}^{n-1} \to [-\frac{1}{2}, \frac{1}{2}]$ such that

$$\nabla b_{\delta,L,M}(z) = 0 \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } z \in \Sigma_{\delta,L,M}, \tag{3-52}$$

$$D^c b_{\delta,L,M}(G) = \int_G f_{\delta,L,M} \, d(D^c v) \quad \text{for every bounded Borel set } G \subset \Sigma_{\delta,L,M}^{(1)}. \tag{3-53}$$

Indeed, let us set $v_{\delta,L,M} = 1_{\Sigma_{\delta,L,M}} v$. Since $v_{\delta,L,M}, b_{\delta,L,M} \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, we can apply Corollary 3.3 to $W = W[v_{\delta,L,M}, b_{\delta,L,M}]$. Since $W[v_{\delta,L,M}, b_{\delta,L,M}] = E \cap (\Sigma_{\delta,L,M} \times \mathbb{R})$, and thus

$$\partial^e E \cap (\Sigma_{\delta,L,M}^{(1)} \times \mathbb{R}) = \partial^e W[v_{\delta,L,M}, b_{\delta,L,M}] \cap (\Sigma_{\delta,L,M}^{(1)} \times \mathbb{R}),$$

we find that, for every Borel set $G \subset \Sigma_{\delta,L,M}^{(1)} \setminus (S_{v_{\delta,L,M}} \cup S_{b_{\delta,L,M}})$,

$$
\begin{aligned}
P(E; G \times \mathbb{R}) &= P(W[v_{\delta,L,M}, b_{\delta,L,M}]; G \times \mathbb{R}) \\
&= \int_G \sqrt{1 + |\nabla(b_{\delta,L,M} + \tfrac{1}{2} v_{\delta,L,M})|^2} + \sqrt{1 + |\nabla(b_{\delta,L,M} - \tfrac{1}{2} v_{\delta,L,M})|^2} \, d\mathcal{H}^{n-1} \\
&\quad + |D^c(b_{\delta,L,M} + \tfrac{1}{2} v_{\delta,L,M})|(G) + |D^c(b_{\delta,L,M} - \tfrac{1}{2} v_{\delta,L,M})|(G).
\end{aligned}
\tag{3-54}
$$

By Lemma 2.3 applied to $v_{\delta,L,M} = 1_{\Sigma_{\delta,L,M}} v$, we find that

$$\nabla v_{\delta,L,M} = 1_{\Sigma_{\delta,L,M}} \nabla v \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1},$$

$$D^c v_{\delta,L,M} = D^c v \llcorner \Sigma_{\delta,L,M}^{(1)}, \quad S_{v_{\delta,L,M}} \cap \Sigma_{\delta,L,M}^{(1)} = S_v \cap \Sigma_{\delta,L,M}^{(1)}.$$

By (3-54), we thus find that

$$
\begin{aligned}
P(E; G \times \mathbb{R}) &= \int_G \sqrt{1 + |\nabla(b_{\delta,L,M} + \tfrac{1}{2} v)|^2} + \sqrt{1 + |\nabla(b_{\delta,L,M} - \tfrac{1}{2} v)|^2} \, d\mathcal{H}^{n-1} \\
&\quad + |D^c(b_{\delta,L,M} + \tfrac{1}{2} v)|(G) + |D^c(b_{\delta,L,M} - \tfrac{1}{2} v)|(G)
\end{aligned}
\tag{3-55}
$$

for every Borel set $G \subset \Sigma_{\delta,L,M}^{(1)} \setminus (S_v \cup S_{b_{\delta,L,M}})$. By Corollary 3.4, for every Borel set $G \subset \mathbb{R}^{n-1}$,

$$P(F[v]; G \times \mathbb{R}) = 2 \int_G \sqrt{1 + |\tfrac{1}{2} \nabla v|^2} \, d\mathcal{H}^{n-1} + \int_{G \cap S_v} [v] \, d\mathcal{H}^{n-2} + |D^c v|(G). \tag{3-56}$$

Taking into account that $P(E; G \times \mathbb{R}) = P(F[v]; G \times \mathbb{R})$ for every Borel set $G \subset \mathbb{R}^{n-1}$, we combine (3-55) and (3-56), together with the convexity of the map $\xi \mapsto \sqrt{1 + |\xi|^2}$, $\xi \in \mathbb{R}^{n-1}$, and (3-49), to find that, if $G \subset \Sigma_{\delta,L,M}^{(1)} \setminus (S_v \cup S_{b_{\delta,L,M}})$, then

$$0 = \int_G \sqrt{1 + |\nabla(b_{\delta,L,M} + \tfrac{1}{2} v)|^2} + \sqrt{1 + |\nabla(b_{\delta,L,M} - \tfrac{1}{2} v)|^2} - 2\sqrt{1 + |\tfrac{1}{2} \nabla v|^2} \, d\mathcal{H}^{n-1}, \tag{3-57}$$

$$0 = |D^c(b_{\delta,L,M} + \tfrac{1}{2} v)|(G) + |D^c(b_{\delta,L,M} - \tfrac{1}{2} v)|(G) - |D^c v|(G). \tag{3-58}$$

Since $\Sigma^{(1)}_{\delta,L,M} \setminus (S_v \cup S_{b_{\delta,L,M}})$ is $\mathcal{H}^{n-1}$-equivalent to $\Sigma_{\delta,L,M}$, by (3-57) and by the strict convexity of $\xi \in \mathbb{R}^{n-1} \mapsto \sqrt{1 + |\xi|^2}$ we obtain (3-52). By applying Lemma 3.7 to

$$\mu = \tfrac{1}{2}D^c v, \quad \nu = D^c b_{\delta,L,M} \llcorner \big(\Sigma^{(1)}_{\delta,L,M} \setminus (S_v \cup S_{b_{\delta,L,M}})\big) = D^c b_{\delta,L,M} \llcorner \Sigma^{(1)}_{\delta,L,M},$$

we prove (3-53). This completes the proof of (3-52) and (3-53).

Step two: We prove (1-18). Let $\delta, L \in I$ and $M \in J_\delta$. Since $b_{\delta,L,M} = 1_{\Sigma_{\delta,L,M}} \tau_M b_\delta$, by Lemma 2.3 we have

$$D^c b_{\delta,L,M} = D^c (\tau_M b_\delta) \llcorner \Sigma^{(1)}_{\delta,L,M}.$$

We combine this fact with (3-53) to find a Borel function $f_{\delta,M} : \mathbb{R}^{n-1} \to [-\tfrac{1}{2}, \tfrac{1}{2}]$ with

$$D^c \tau_M b_\delta(G) = \int_G f_{\delta,M} \, d(D^c v) \quad \text{for every bounded Borel set } G \subset \Sigma^{(1)}_{\delta,L,M}.$$

As a consequence, the Radon measures $D^c \tau_M b_\delta$ and $f_{\delta,M} D^c v$ coincide on every bounded Borel set contained in

$$\begin{aligned}
\bigcup_{L \in I} \Sigma^{(1)}_{\delta,L,M} &= \bigcup_{L \in I} \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)} \cap \{v < L\}^{(1)} \\
&= \big(\{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}\big) \cap \bigcup_{L \in I} \{v < L\}^{(1)} \\
&= \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)} \cap \{v^\vee < \infty\},
\end{aligned}$$

where in the last identity we have used (2-8). Since $\mathcal{H}^{n-2}(\{v^\vee = \infty\}) = 0$ by [Federer 1969, 4.5.9(3)], the set $\{v^\vee = \infty\}$ is negligible with respect to both $|D^c \tau_M b_\delta|$ and $|D^c v|$. We have thus proved that, for every bounded Borel set $G \subset \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}$,

$$D^c (\tau_M b_\delta)(G) = \int_G f_{\delta,M} \, d(D^c v). \tag{3-59}$$

Since for every $M' > M$ and $\delta' < \delta$ we have that $\tau_M b_\delta = \tau_{M'} b_{\delta'}$ on $\{v > \delta\} \cap \{|b_E| < M\}$, by Lemma 2.2 we obtain that

$$D^c (\tau_M b_\delta) \llcorner \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)} = D^c (\tau_{M'} b_{\delta'}) \llcorner \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)},$$

and therefore (3-59) can be rewritten with a function f independent of M and δ; thus,

$$D^c (\tau_M b_\delta)(G) = \int_G f \, d(D^c v) \tag{3-60}$$

for every bounded Borel set $G \subset \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}$. We next note that, if $\delta \in I$ and $M \in J_\delta$, then

$$\tau_M b_\delta = M 1_{\{b_\delta \geq M\}} - M 1_{\{b_\delta \leq -M\}} + 1_{\{|b_\delta| < M\} \cap \{v > \delta\}} \tau_M b_\delta \quad \text{on } \mathbb{R}^{n-1}$$

is an identity between BV functions. By [Ambrosio et al. 2000, Example 3.97] we thus find

$$\begin{aligned}
D^c \tau_M b_\delta &= D^c (1_{\{|b_\delta| < M\} \cap \{v > \delta\}} \tau_M b_\delta) = 1_{(\{|b_\delta| < M\} \cap \{v > \delta\})^{(1)}} D^c (\tau_M b_\delta) \\
&= D^c (\tau_M b_\delta) \llcorner \big(\{|b_\delta| < M\}^{(1)} \cap \{v > \delta\}^{(1)}\big).
\end{aligned} \tag{3-61}$$

Since, by (3-61), the measure $D^c(\tau_M b_\delta)$ is concentrated on $\{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}$, we deduce from (3-60) that, for every bounded Borel set $G \subset \mathbb{R}^{n-1}$,

$$D^c(\tau_M b_\delta)(G) = D^c(\tau_M b_\delta)\big(G \cap \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}\big) = \int_{G \cap \{v > \delta\}^{(1)} \cap \{|b_E| < M\}^{(1)}} f \, d(D^c v),$$

which proves (1-18).

Step three: We prove (1-16). Let $\delta, L \in I$ and $M \in J_\delta$. Since $b_{\delta,L,M} = b_E$ on $\Sigma_{\delta,L,M}$, by (3-52) and by (2-12) we find that $\nabla b_E = 0$ $\mathcal{H}^{n-1}$-a.e. on $\Sigma_{\delta,L,M}$. By taking a union first over $M \in J_\delta$, and then over $\delta, L \in I$, we find that $\nabla b_E = 0$ $\mathcal{H}^{n-1}$-a.e. on $\{v > 0\}$. At the same time, $b_E = 0$ on $\{v = 0\}$ by definition, and thus, again by (2-12), we have $\nabla b_E = 0$ $\mathcal{H}^{n-1}$-a.e. on $\{v = 0\}$. This completes the proof of (1-16).

Step four: We prove (1-17). We fix $\delta, L \in I$ and define $\Sigma_{\delta,L} = \{\delta < v < L\}$, $b_{\delta,L} = 1_{\Sigma_{\delta,L}} b_E$, and $v_{\delta,L} = 1_{\Sigma_{\delta,L}} v$. Since $\Sigma_{\delta,L}$ is a set of finite perimeter, $b_{\delta,L} \in GBV(\mathbb{R}^{n-1})$, while, by construction, $v_{\delta,L} \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$. We are in position to apply Corollary 3.3 to obtain a formula for the perimeter of $W[v_{\delta,L}, b_{\delta,L}]$ relative to cylinders $G \times \mathbb{R}$ for Borel sets $G \subset \mathbb{R}^{n-1}$. In particular, if $G \subset \Sigma_{\delta,L}^{(1)} \cap (S_{v_{\delta,L}} \cup S_{b_{\delta,L}})$, then

$$P(E; G \times \mathbb{R}) = P(W[v_{\delta,L}, b_{\delta,L}]; G \times \mathbb{R}) = \int_G \min\{v_{\delta,L}^\vee + v_{\delta,L}^\wedge, \max\{[v_{\delta,L}], 2[b_{\delta,L}]\}\} \, d\mathcal{H}^{n-2}.$$

Since, by (2-10), $\Sigma_{\delta,L}^{(1)} \cap S_{v_{\delta,L}} = \Sigma_{\delta,L}^{(1)} \cap S_v$ with $v_{\delta,L}^\vee = v^\vee$, $v_{\delta,L}^\wedge = v^\wedge$, and $[v_{\delta,L}] = [v]$ on $\Sigma_{\delta,L}^{(1)}$, we have

$$P(E; G \times \mathbb{R}) = \int_G \min\{v^\vee + v^\wedge, \max\{[v], 2[b_{\delta,L}]\}\} \, d\mathcal{H}^{n-2}$$

whenever $G \subset \Sigma_{\delta,L}^{(1)} \cap (S_v \cup S_{b_{\delta,L}})$. Since $P(E; G \times \mathbb{R}) = P(F[v]; G \times \mathbb{R})$, by (3-56),

$$\min\{v^\vee + v^\wedge, \max\{[v], 2[b_{\delta,L}]\}\} = [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } (S_{b_{\delta,L}} \cup S_v) \cap \Sigma_{\delta,L}^{(1)}.$$

Since $v^\wedge \geq \delta$ on $\Sigma_{\delta,L}^{(1)}$, we deduce that $v^\vee + v^\wedge > [v]$ on $\Sigma_{\delta,L}^{(1)}$, and thus the above condition immediately implies that

$$2[b_{\delta,L}] \leq [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } (S_{b_{\delta,L}} \cup S_v) \cap \Sigma_{\delta,L}^{(1)}.$$

In particular, $S_{b_{\delta,L}} \cap \Sigma_{\delta,L}^{(1)} \subset_{\mathcal{H}^{n-2}} S_v$, and we have proved

$$2[b_{\delta,L}] \leq [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } \Sigma_{\delta,L}^{(1)}.$$

By (2-10), $[b_{\delta,L}] = [b_E]$ on $\Sigma_{\delta,L}^{(1)}$. By taking the union of $\Sigma_{\delta,L}^{(1)}$ on $\delta, L \in I$, and using (2-8) and (2-9), we find that

$$2[b_E] \leq [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } \{v^\wedge > 0\} \cup \{v^\vee < \infty\}.$$

Since, as noted above, $\{v^\vee = \infty\}$ is $\mathcal{H}^{n-2}$-negligible, we have proved (1-17). $\qquad \square$

3D. *Characterization of equality cases, part two.*
We now complete the proof of Theorem 1.9, by showing that if a v-distributed set of finite perimeter E satisfies (1-15), (1-16), (1-17), and (1-18), then $E \in \mathcal{M}(v)$. The following proposition will play a crucial role.

Proposition 3.8. *If* $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$, $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, *and* E *is a* v-*distributed set of finite perimeter with segments as sections, then*

$$P(E; \{v^\wedge = 0\} \times \mathbb{R}) = P(F[v]; \{v^\wedge = 0\} \times \mathbb{R}) = \int_{\{v^\wedge = 0\}} v^\vee \, d\mathcal{H}^{n-2}. \tag{3-62}$$

Remark 3.9. With Proposition 3.8, one can actually go back to Corollary 3.3 and obtain a formula for $P(E; G \times \mathbb{R})$ in terms of v and b_E whenever E is a v-distributed set of finite perimeter with segments as sections. Since such a formula may be of independent interest, we have included its proof in Appendix B.

Proof of Proposition 3.8. Let $I = \{t > 0 : \{v > t\}$ and $\{v < t\}$ are of finite perimeter$\}$, so that we have, as usual, $\mathcal{H}^1((0, \infty) \setminus I) = 0$. Since

$$\int_0^\infty P(\{v > t\}) \, dt = \int_0^\infty P(\{v < t\}) \, dt = |Dv|(\mathbb{R}^{n-1}) < \infty,$$

we can find two sequences $\{\delta_h\}_{h \in \mathbb{N}}, \{L_h\}_{h \in \mathbb{N}} \subset I$ such that

$$\lim_{h \to \infty} \delta_h = 0, \qquad\qquad \lim_{h \to \infty} \delta_h P(\{v > \delta_h\}) = 0, \tag{3-63}$$

$$\lim_{h \to \infty} L_h = \infty, \qquad\qquad \lim_{h \to \infty} L_h P(\{v < L_h\}) = 0. \tag{3-64}$$

Let us set $\Sigma_h = \{L_h > v > \delta_h\}$ and $E_h = E \cap (\Sigma_h \times \mathbb{R})$. Note that E_h is, trivially, a set of locally finite perimeter. Now, E_h locally converges to E as $h \to \infty$, and also $P(E_h; \Sigma_h^{(0)} \times \mathbb{R}) = 0$ and $\partial^e E_h \cap (\Sigma_h^{(1)} \times \mathbb{R}) = \partial^e E \cap (\Sigma_h^{(1)} \times \mathbb{R})$, so we have

$$P(E) \leq \liminf_{h \to \infty} P(E_h) = \liminf_{h \to \infty} P(E; \Sigma_h^{(1)} \times \mathbb{R}) + P(E_h; \partial^e \Sigma_h \times \mathbb{R}). \tag{3-65}$$

By (2-8) and (2-9),

$$\lim_{h \to \infty} 1_{\Sigma_h^{(1)}}(z) = 1_{\{v^\wedge > 0\} \cap \{v^\vee < \infty\}}(z) \quad \text{for all } z \in \mathbb{R}^{n-1},$$

so that, by dominated convergence and thanks to the fact that E has finite perimeter,

$$\lim_{h \to \infty} P(E; \Sigma_h^{(1)} \times \mathbb{R}) = P(E; (\{v^\vee > 0\} \cap \{v^\vee < \infty\}) \times \mathbb{R}) = P(E; \{v^\wedge > 0\} \times \mathbb{R}).$$

(In the last identity we have first used [Federer 1969, 4.5.9(3)] to infer that $\mathcal{H}^{n-2}(\{v^\vee = \infty\}) = 0$, and then [Federer 1969, 2.10.45] to conclude that $\mathcal{H}^{n-1}(\{v^\vee = \infty\} \times \mathbb{R}) = 0$.) Hence, by (3-65),

$$P(E; \{v^\wedge = 0\} \times \mathbb{R}) \leq \liminf_{h \to \infty} P(E_h; \partial^e \Sigma_h \times \mathbb{R}). \tag{3-66}$$

Since $\delta_h, L_h \in I$, we have $v_h = 1_{\Sigma_h} v \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ and $a_h = 1_{\Sigma_h} b_E \in GBV(\mathbb{R}^{n-1})$ (indeed, $a_h = 1_{\{v < L_h\}} b_{\delta_h}$, where $b_{\delta_h} = 1_{\{v > \delta_h\}} b_E \in GBV(\mathbb{R}^{n-1})$, thanks to Theorem 1.7). Since $E_h = W[v_h, a_h]$ according to (3-24), we can apply (3-25) in Corollary 3.3 to $G = \partial^e \Sigma_h$ to find that

$$P(E_h; \partial^e \Sigma_h \times \mathbb{R}) = \int_{\partial^e \Sigma_h \cap (S_{v_h} \cup S_{a_h})} \min\{v_h^\vee + v_h^\wedge, \max\{[v_h], 2[a_h]\}\} \, d\mathcal{H}^{n-2}. \tag{3-67}$$

Note that, since $\partial^e \Sigma_h$ is countably $\mathcal{H}^{n-2}$-rectifiable, we are only interested in the "jump" contribution in (3-25). Let us now set

$$K_h^1 = \partial^e \Sigma_h \cap \partial^e \{v > \delta_h\}, \quad K_h^2 = \partial^e \Sigma_h \setminus \partial^e \{v > \delta_h\} \subset \partial^e \{v < L_h\}.$$

The key observation to exploit (3-67) is that, as one can check with standard arguments,

$$v_h^\vee = v^\vee \geq \delta_h \geq v^\wedge \quad \text{and} \quad v_h^\wedge = 0 \quad \mathcal{H}^{n-2}\text{-a.e. on } K_h^1, \tag{3-68}$$

$$v^\vee \geq L_h \geq v^\wedge = v_h^\vee \quad \text{and} \quad v_h^\wedge = 0 \quad \mathcal{H}^{n-2}\text{-a.e. on } K_h^2. \tag{3-69}$$

For example, in order to prove (3-69), we argue as follows. First, we note that we always have $v^\vee \geq L_h \geq v^\wedge$ and $v_h^\wedge = 0$ on $\partial^e \{v < L_h\}$. In particular, $\tilde{v} = L_h$ on $S_v^c \cap \partial^e \{v < L_h\}$, and this immediately implies $v_h^\vee = L_h$ on $S_v^c \cap \partial^e \{v < L_h\}$. By noting that $v_h = 1_{\Sigma_h} v$ with $\Sigma_h \subset \{v < L_h\}$, one checks that $v^\wedge = v_h^\vee$ $\mathcal{H}^{n-2}$-a.e. on $J_v \cap \partial^* \{v < L_h\}$. By (3-68) and (3-69), we have

$$\min\{v_h^\vee + v_h^\wedge, \max\{[v_h], 2[a_h]\}\} = v^\vee \quad \mathcal{H}^{n-2}\text{-a.e. on } K_h^1, \tag{3-70}$$

$$\min\{v_h^\vee + v_h^\wedge, \max\{[v_h], 2[a_h]\}\} = v^\wedge \quad \mathcal{H}^{n-2}\text{-a.e. on } K_h^2, \tag{3-71}$$

so that, by (3-67) and since $K_h^1 \subset_{\mathcal{H}^{n-2}} S_{v_h}$ — which again follows from (3-68) — we find

$$P(E_h; \partial^e \Sigma_h \times \mathbb{R}) \leq \int_{K_h^1} v^\vee \, d\mathcal{H}^{n-2} + \int_{K_h^2} v^\wedge \, d\mathcal{H}^{n-2}. \tag{3-72}$$

By (3-69) and (3-64), we have

$$\limsup_{h \to \infty} \int_{K_h^2} v^\wedge \, d\mathcal{H}^{n-2} \leq \limsup_{h \to \infty} L_h \mathcal{H}^{n-2}(K_h^2) \leq \limsup_{h \to \infty} L_h P(\{v < L_h\}) = 0. \tag{3-73}$$

We are now going to prove that

$$\lim_{h \to \infty} \int_{\partial^e \{v > \delta_h\}} v^\vee \, d\mathcal{H}^{n-2} = \int_{\{v^\wedge = 0\}} v^\vee \, d\mathcal{H}^{n-2}. \tag{3-74}$$

This will be useful in the estimate of the right-hand side of (3-67) because $K_h^1 \subset \partial^e \{v > \delta_h\}$. Since $\{v^\wedge = 0\} \cap \partial^e \{v > \delta_h\} = \{v^\wedge = 0\} \cap S_v \cap \partial^e \{v > \delta_h\} = \{v^\wedge = 0\} \cap \{[v] \geq \delta_h\}$, we have that, monotonically as $h \to \infty$,

$$v^\vee 1_{\{v^\wedge = 0\} \cap \partial^e \{v > \delta_h\}} \to v^\vee 1_{\{v^\wedge = 0\} \cap S_v} \quad \text{pointwise on } \mathbb{R}^{n-1}.$$

Hence,

$$\lim_{h \to \infty} \int_{\{v^\wedge = 0\} \cap \partial^e \{v > \delta_h\}} v^\vee \, d\mathcal{H}^{n-2} = \int_{\{v^\wedge = 0\} \cap S_v} v^\vee \, d\mathcal{H}^{n-2} = \int_{\{v^\wedge = 0\}} v^\vee \, d\mathcal{H}^{n-2}. \tag{3-75}$$

We now claim that

$$\lim_{h \to \infty} \int_{\{v^\wedge > 0\} \cap \partial^e \{v > \delta_h\}} v^\vee \, d\mathcal{H}^{n-2} = 0. \tag{3-76}$$

Indeed, since $v^\vee = v^\wedge = \delta_h$ on $S_v^c \cap \partial^e \{v > \delta_h\}$, we find that

$$\int_{S_v^c \cap \{v^\wedge > 0\} \cap \partial^e \{v > \delta_h\}} v^\vee \, d\mathcal{H}^{n-2} \leq \delta_h \mathcal{H}^{n-2}(\partial^e \{v > \delta_h\}) = \delta_h P(\{v > \delta_h\}),$$

so that, by (3-63),

$$
\begin{aligned}
\limsup_{h\to\infty} \int_{\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} v^\vee \, d\mathcal{H}^{n-2} &= \limsup_{h\to\infty} \int_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} v^\vee \, d\mathcal{H}^{n-2} \\
&= \limsup_{h\to\infty} \int_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} [v] + v^\wedge \, d\mathcal{H}^{n-2} \\
&\leq \limsup_{h\to\infty} \int_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} [v] \, d\mathcal{H}^{n-2} + \delta_h \mathcal{H}^{n-2}(\partial^e\{v>\delta_h\}) \\
&= \limsup_{h\to\infty} \int_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} [v] \, d\mathcal{H}^{n-2}, \qquad (3\text{-}77)
\end{aligned}
$$

where the inequality follows by (3-68), and the last equality is by (3-63). Now, if $z \in \{v^\wedge > 0\}$, then $z \in \{v > \delta\}^{(1)}$ for every $\delta < v^\wedge(z)$, so that

$$
1_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} \to 0 \quad \text{pointwise on } \mathbb{R}^{n-1}
$$

as $h \to \infty$. Since $[v] \in L^1(\mathcal{H}^{n-2} \llcorner S_v)$, by dominated convergence we find

$$
\lim_{h\to\infty} \int_{S_v\cap\{v^\wedge>0\}\cap\partial^e\{v>\delta_h\}} [v] \, d\mathcal{H}^{n-2} = 0. \qquad (3\text{-}78)
$$

By combining (3-77) and (3-78), we obtain (3-76). By (3-75) and (3-76), we deduce (3-74). From $K_h^1 \subset \partial^e\{v > \delta_h\}$, (3-72), (3-73), and (3-74), we deduce that

$$
\limsup_{h\to\infty} P(E_h; \partial^e\Sigma_h \times \mathbb{R}) \leq \int_{\{v^\wedge=0\}} v^\vee \, d\mathcal{H}^{n-2}.
$$

By combining this last inequality with (3-66), we find

$$
P(E; \{v^\wedge=0\} \times \mathbb{R}) \leq \int_{\{v^\wedge=0\}} v^\vee \, d\mathcal{H}^{n-2} = P(F[v]; \{v^\wedge=0\} \times \mathbb{R}) \leq P(E; \{v^\wedge=0\} \times \mathbb{R}),
$$

where the equality follows by (3-29), and the final inequality is, of course, (1-1). This completes the proof of (3-62). $\qquad\square$

Remark 3.10. Let $v \in BV(\mathbb{R}^{n-1}; [0,\infty))$ with $\mathcal{H}^{n-1}(\{v>0\}) < \infty$, and let E be a v-distributed set with segments as sections. Then, E is of finite perimeter if and only if $\sup_{h\in\mathbb{N}} P(E_h) < \infty$, where

$$
E_h = E \cap (\Sigma_h \times \mathbb{R}), \quad \Sigma_h = \{L_h > v > \delta_h\},
$$

and $\{\delta_h\}_{h\in\mathbb{N}}, \{L_h\}_{h\in\mathbb{N}} \subset (0,\infty)$ are such that

$$
\begin{aligned}
\lim_{h\to\infty} \delta_h = 0, &\qquad\qquad \lim_{h\to\infty} \delta_h P(\{v>\delta_h\}) = 0, \\
\lim_{h\to\infty} L_h = \infty, &\qquad\qquad \lim_{h\to\infty} L_h P(\{v<L_h\}) = 0.
\end{aligned}
$$

The fact that $P(E) < \infty$ implies $\sup_{h\in\mathbb{N}} P(E_h) < \infty$ is implicit in the proof of Proposition 3.8. Conversely, if $\{E_h\}_{h\in\mathbb{N}}$ is defined as above, then $E_h \to E$ as $h \to \infty$, and thus $\sup_{h\in\mathbb{N}} P(E_h) < \infty$ implies $P(E) < \infty$ by lower semicontinuity of perimeter.

Lemma 3.11. *If $v \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, $b : \mathbb{R}^{n-1} \to \mathbb{R}$ is such that $\tau_M b \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ for a.e. $M > 0$, and μ is an $\mathbb{R}^{n-1}$-valued Radon measure such that*

$$\lim_{M \to \infty} |\mu - D^c \tau_M b|(G) = 0 \quad \text{for every bounded Borel set } G \subseteq \mathbb{R}^{n-1}, \tag{3-79}$$

then

$$|D^c(b+v)|(G) \le |\mu + D^c v|(G) \quad \text{for every Borel set } G \subseteq \mathbb{R}^{n-1}. \tag{3-80}$$

Proof. Let us assume that $|v| \le L \ \mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$. If $f \in BV(\mathbb{R}^{n-1})$, then

$$\tau_M f = M 1_{\{f > M\}} - M 1_{\{f < -M\}} + 1_{\{|f| < M\}} \tau_M f \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$$

for every M such that $\{f > M\}$ and $\{f < -M\}$ are of finite perimeter, and thus, by [Ambrosio et al. 2000, Example 3.97],

$$D^c \tau_M f = D^c(1_{\{|f| < M\}} \tau_M f) = 1_{\{|f| < M\}^{(1)}} D^c(\tau_M f) = D^c(\tau_M f) \llcorner \{|f| < M\}^{(1)};$$

in particular,

$$|D^c \tau_M f| = |D^c f| \llcorner \{|f| < M\}^{(1)} \le |D^c f|. \tag{3-81}$$

From the equality $\tau_M(\tau_{M+L}(b) + v) = \tau_M(b+v)$ and from (3-81) applied with $f = \tau_{M+L}(b) + v$ it follows that, for every Borel set $G \subseteq \mathbb{R}^{n-1}$,

$$\left| D^c(\tau_M(b+v)) \right|(G) = \left| D^c\bigl(\tau_M(\tau_{M+L}(b) + v)\bigr) \right|(G) \le \left| D^c(\tau_{M+L}(b) + v) \right|(G). \tag{3-82}$$

By (3-79),

$$\lim_{M \to \infty} \left| D^c(\tau_{M+L}(b) + v) \right|(G) = |\mu + D^c v|(G).$$

We let $M \to \infty$ in (3-82), and by definition of $|D^c(b+v)|$ we obtain (3-80).   □

Proof of Theorem 1.9 (sufficient conditions). Let E be a v-distributed set of finite perimeter satisfying (1-15), (1-16), (1-17), and (1-18). Let I and J_δ be defined as in (3-50) and (3-51). If $\delta, S \in I$ and we set $b_{\delta,S} = 1_{\{\delta < v < S\}} b_E = 1_{\{\delta < v < S\}} b_\delta$, then, for every $M \in J_\delta$, we have $\tau_M b_\delta \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ (see the end of step one in the proof of Theorem 1.7), and so we obtain that $\tau_M b_{\delta,S} \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$. Let us consider the $\mathbb{R}^{n-1}$-valued Radon measure $\mu_{\delta,S}$ on $\mathbb{R}^{n-1}$ defined for every bounded Borel set $G \subset \mathbb{R}^{n-1}$ by

$$\mu_{\delta,S}(G) = \int_{G \cap \{\delta < v < S\}^{(1)} \cap \{|b_E|^\vee < \infty\}} f \, d D^c v$$

Since $\tau_M b_{\delta,S} = 1_{\{v < S\}} \tau_M b_\delta$, by Lemma 2.3 we have $D^c[\tau_M b_{\delta,S}] = 1_{\{v < S\}^{(1)}} D^c[\tau_M b_\delta]$, and thus, for every Borel set $G \subset \mathbb{R}^{n-1}$,

$$\lim_{M \to \infty} |\mu_{\delta,S} - D^c[\tau_M b_{\delta,S}]|(G) = \lim_{M \to \infty} |\mu_{\delta,S} - D^c[\tau_M b_\delta]|(G \cap \{v < S\}^{(1)})$$

$$\le \lim_{M \to \infty} \int_{G \cap \{\delta < v < S\}^{(1)} \cap [\{|b_E|^\vee < \infty\} \setminus \{|b_E| < M\}^{(1)}]} |f| \, d|D^c v| = 0, \tag{3-83}$$

where the inequality follows by (1-18), and the last equality follows from the fact that $\{\{|b_E| < M\}^{(1)}\}_{M \in I}$ is an increasing family of sets whose union is $\{|b_E|^\vee < \infty\}$. By applying Lemma 3.11 to $b_{\delta,S}$ and $\pm \frac{1}{2} v_{\delta,S}$

(with $v_{\delta,S} = 1_{\{\delta < v < S\}}v$), and Lemma 3.7 to $\mu_{\delta,S}$ and $\pm\frac{1}{2}D^c v_{\delta,S}$ and recalling (1-18), we find that, for every bounded Borel set $G \subset \mathbb{R}^{n-1}$,

$$|D^c(b_{\delta,S} + \tfrac{1}{2}v_{\delta,S})|(G) + |D^c(b_{\delta,S} - \tfrac{1}{2}v_{\delta,S})|(G) \le |\mu_{\delta,S} + \tfrac{1}{2}D^c v_{\delta,S}|(G) + |\mu_{\delta,S} - \tfrac{1}{2}D^c v_{\delta,S}|(G)$$
$$= |D^c v_{\delta,S}|(G). \tag{3-84}$$

Since $b_{\delta,S} \in GBV(\mathbb{R}^{n-1})$ and $v_{\delta,S} \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$, if $W = W[v_{\delta,S}, b_{\delta,S}]$, then we can compute $P(W; G \times \mathbb{R})$ for every Borel set $G \subset \mathbb{R}^{n-1}$ by Corollary 3.3. In particular, if $G \subset \{\delta < v < S\}^{(1)}$, then by $E \cap (\{\delta < v < S\} \times \mathbb{R}) = W \cap (\{\delta < v < S\} \times \mathbb{R})$ we find that

$$P(E; G \times \mathbb{R}) = P(W; G \times \mathbb{R})$$
$$= \int_G \sqrt{1 + |\nabla(b_{\delta,S} + \tfrac{1}{2}v_{\delta,S})|^2} + \sqrt{1 + |\nabla(b_{\delta,S} - \tfrac{1}{2}v_{\delta,S})|^2} \, d\mathcal{H}^{n-1}$$
$$+ \int_{G \cap (S_{v_{\delta,S}} \cup S_{b_{\delta,S}})} \min\{v^\vee_{\delta,S} + v^\wedge_{\delta,S}, \max\{[v_{\delta,S}], 2[b_{\delta,S}]\}\} \, d\mathcal{H}^{n-2}$$
$$+ |D^c(b_{\delta,S} + \tfrac{1}{2}v_{\delta,S})|(G) + |D^c(b_{\delta,S} - \tfrac{1}{2}v_{\delta,S})|(G). \tag{3-85}$$

We can also compute $P(F[v_{\delta,S}]; G \times \mathbb{R})$ using Corollary 3.4. Since

$$F[v] \cap (\{\delta < v < S\} \times \mathbb{R}) = F[v_{\delta,S}] \cap (\{\delta < v < S\} \times \mathbb{R}),$$

we conclude that

$$P(F; G \times \mathbb{R}) = P(F[v_{\delta,S}]; G \times \mathbb{R})$$
$$= 2\int_G \sqrt{1 + |\tfrac{1}{2}\nabla v_{\delta,S}|^2} \, d\mathcal{H}^{n-1} + \int_{G \cap S_{v_{\delta,S}}} [v_{\delta,S}] \, d\mathcal{H}^{n-2} + |D^c v_{\delta,S}|(G). \tag{3-86}$$

From (1-16) and (1-17) we deduce that (applying (2-10) and (2-12) to b_E and v)

$$\nabla b_{\delta,S}(z) = \nabla b_E = 0 \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } z \in \{\delta < v < S\}, \tag{3-87}$$
$$2[b_{\delta,S}] = 2[b_E] \le [v] = [v_{\delta,S}] \quad \mathcal{H}^{n-2}\text{-a.e. on } \{\delta < v < S\}^{(1)}. \tag{3-88}$$

Substituting (3-87), (3-88), and (3-84) into the first, second, and third parts of (3-85) respectively, we find that

$$P(E; \{\delta < v < S\}^{(1)} \times \mathbb{R}) \le P(F; \{\delta < v < S\}^{(1)} \times \mathbb{R}), \tag{3-89}$$

where, in fact, equality holds thanks to (1-1). By (2-9) it follows that

$$\bigcup_{M \in I} \{v < M\}^{(1)} = \{v^\vee < \infty\} =_{\mathcal{H}^{n-2}} \mathbb{R}^{n-1}, \tag{3-90}$$

as $\mathcal{H}^{n-2}(\{v^\vee = \infty\}) = 0$ by [Federer 1969, 4.5.9(3)]. By taking a union over $\delta_h \in I$ and $S_h \in I$ such that $\delta_h \to 0$ and $S_h \to \infty$ as $h \to \infty$, we deduce from (3-89), (3-48), and (3-90) that

$$P(E; \{v^\wedge > 0\} \times \mathbb{R}) = P(F; \{v^\wedge > 0\} \times \mathbb{R}).$$

By Proposition 3.8, $P(E; \{v^\wedge = 0\} \times \mathbb{R}) = P(F; \{v^\wedge = 0\} \times \mathbb{R})$, and thus $P(E) = P(F)$, as required. $\square$

3E. _Equality cases by countably many vertical translations._ We finally address the problem of characterizing the situation when equality cases are necessarily obtained by countably many vertical translations of parts of $F[v]$; see (1-22). In particular, we want to show this situation is characterized by the assumptions that $v \in SBV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$ and S_v is locally $\mathcal{H}^{n-2}$-rectifiable. We shall need:

Theorem 3.12. _Let_ $u : \mathbb{R}^{n-1} \to \mathbb{R}$ _be Lebesgue measurable. The following are equivalent:_

(i) $u \in GBV(\mathbb{R}^{n-1})$ _with_ $|D^c u| = 0$, $\nabla u = 0$ $\mathcal{H}^{n-1}$-_a.e. on_ $\mathbb{R}^{n-1}$, _and_ S_u _locally_ $\mathcal{H}^{n-2}$-_finite._

(ii) _There exist an at most countable set_ I, $\{c_h\}_{h \in I} \subset \mathbb{R}$, _and a partition_ $\{G_h\}_{h \in I}$ _of_ $\mathbb{R}^{n-1}$ _into Borel sets such that_

$$u = \sum_{h \in I} c_h 1_{G_h} \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1} \tag{3-91}$$

and $\sum_{h \in I} P(G_h \cap B_R) < \infty$ _for every_ $R > 0$.

Moreover, if we assume that $c_h \neq c_k$ _for_ $h \neq k \in I$ _then, when_ (i) _and_ (ii) _hold,_

$$S_u \subset_{\mathcal{H}^{n-2}} \bigcup_{h \neq k \in I} \partial^e G_h \cap \partial^e G_k \tag{3-92}$$

with $[u] = |c_h - c_k|$ $\mathcal{H}^{n-2}$-_a.e. on_ $\partial^e G_h \cap \partial^e G_k$. _In particular,_

$$\sum_{h \in I} P(G_h; B_R) = 2\mathcal{H}^{n-2}(S_u \cap B_R) \quad \text{for all } R > 0.$$

Proof of Theorem 3.12. _Step one:_ We recall that, by [Ambrosio et al. 2000, Definitions 4.16 and 4.21, Theorem 4.23], for every open set Ω and $u \in L^\infty(\Omega)$, the following two conditions are equivalent:

(j) There exist an at most countable set I, $\{c_h\}_{h \in I} \subset \mathbb{R}$, and a partition $\{G_h\}_{h \in I}$ of Ω such that $\sum_{h \in I} P(G_h; \Omega) < \infty$ and

$$u = \sum_{h \in I} c_h 1_{G_h} \quad \mathcal{H}^{n-1}\text{-a.e. on } \Omega. \tag{3-93}$$

(jj) $u \in BV_{\text{loc}}(\Omega)$, $Du = Du \llcorner S_u$, and $\mathcal{H}^{n-2}(S_u \cap \Omega) < \infty$.

When these hold, we have $2\mathcal{H}^{n-2}(S_u \cap \Omega) = \sum_{h \in I} P(G_h; \Omega)$.

Step two: Let us prove that (i) implies (ii). Let $u \in GBV(\mathbb{R}^{n-1})$ with $|D^c u| = 0$, $\nabla u = 0$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, and S_u locally $\mathcal{H}^{n-2}$-finite. For every $R, M > 0$, we have, by the definition of GBV, that $\tau_M u \in BV(B_R)$. Moreover, $|D^c \tau_M u| = 0$, $\nabla \tau_M u = 0$, and $S_{\tau_M u} \cap B_R \subset B_R \cap S_u$ is $\mathcal{H}^{n-2}$-finite. By step one, there exist an at most countable set $I_{R,M}$, $\{c_{R,M,h}\}_{h \in I_{R,M}} \subset \mathbb{R}$, and a partition $\{G_{R,M,h}\}_{h \in I_{R,M}}$ of B_R into sets of finite perimeter such that $\sum_{h \in I_{R,M}} P(G_{R,M,h}; B_R) < \infty$ and

$$\tau_M u = \sum_{h \in I_{R,M}} c_{R,M,h} 1_{G_{R,M,h}} \quad \mathcal{H}^{n-1}\text{-a.e. on } B_R.$$

By a simple monotonicity argument we find (3-91). By (3-91), if we set $J_M = \{h \in \mathbb{N} : |c_h| \leq M\}$ then, $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$,

$$\tau_M u = M 1_{\{u > M\} \cap B_R} - M 1_{\{u < -M\} \cap B_R} + \sum_{h \in J_M} c_h 1_{G_h \cap B_R} \quad \mathcal{H}^{n-1}\text{-a.e. on } B_R. \tag{3-94}$$

By step one,

$$P(\{u > M\}; B_R) + P(\{u < -M\}; B_R) + \sum_{h \in J_M} P(G_h; B_R) = 2\mathcal{H}^{n-2}(S_{\tau_M u} \cap B_R).$$

Thus,

$$\sum_{h \in J_M} P(G_h; B_R) \le 2\mathcal{H}^{n-2}(S_{\tau_M u} \cap B_R) \le 2\mathcal{H}^{n-2}(S_u \cap B_R).$$

Since $\bigcup_{M>0} J_M = I$, letting $M \to \infty$ we find that $\sum_{h \in I} P(G_h; B_R) < \infty$, which clearly implies $\sum_{h \in I} P(G_h \cap B_R) < \infty$.

Step three: We prove that (ii) implies (i). We easily see that, for every R, $M > 0$, $\tau_M u$ satisfies the assumptions (jj) in step one in B_R. Thus, $\tau_M u \in BV(B_R)$ with $D\tau_M u = D\tau_M u \llcorner S_{\tau_M u}$ in B_R, and

$$2\mathcal{H}^{n-2}(S_{\tau_M u} \cap B_R) = \sum_{h \in J_M} P(G_h; B_R) \le \sum_{h \in I} P(G_h \cap B_R) < \infty,$$

where, as before, $J_M = \{h \in \mathbb{N} : |c_h| \le M\}$. This shows that $u \in GBV(\mathbb{R}^{n-1})$ with $|D^c u| = 0$ and $\nabla u = 0$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$. Since $\bigcup_{M>0} S_{\tau_M u} = S_u$, this immediately implies that S_u is locally $\mathcal{H}^{n-2}$-finite.

Step four: We now complete the proof of the theorem. Since $\{G_h\}_{h \in I}$ is an at most countable Borel partition of $\mathbb{R}^{n-1}$ with $\sum_{h \in \mathbb{N}} P(G_h \cap B_R) < \infty$, we have that

$$\mathbb{R}^{n-1} =_{\mathcal{H}^{n-2}} \bigcup_{h \in I} G_h^{(1)} \cup \bigcup_{h \neq k \in I} \partial^e G_h \cap \partial^e G_k;$$

compare with [Ambrosio et al. 2000, Theorem 4.17]. Since $S_u \cap G_h^{(1)} = \varnothing$ for every $h \in I$, this proves (3-92). If we now exploit the fact that, for every $h \neq k \in I$ with $c_h \neq c_k$, G_h and G_k are disjoint sets of locally finite perimeter, then by a blow-up argument we easily see that $[u] = |c_h - c_k|$ $\mathcal{H}^{n-2}$-a.e. on $\partial^e G_h \cap \partial^e G_k$, as required. This completes the proof of theorem. $\qquad\square$

Proof of Theorem 1.13. Step one: We prove that, if $E \in \mathcal{M}(v)$, then there exist a finite or countable set I, $\{c_h\}_{h \in I} \subset \mathbb{R}$, and $\{G_h\}_{h \in I}$ a v-admissible partition of $\{v > 0\}$, such that $b_E = \sum_{h \in I} c_h 1_{G_h}$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$ (so that E satisfies (1-22); see Remark 1.31), $|D^c b_E|^+ = 0$, and $2[b_E] \le [v]$ $\mathcal{H}^{n-2}$-a.e. on $\{v^\wedge > 0\}$. The last two properties of b_E follow immediately from Theorem 1.9 since $D^c v = 0$. We now prove that $b_E = \sum_{h \in I} c_h 1_{G_h}$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$. Let $\delta > 0$ be such that $\{v > \delta\}$ is a set of finite perimeter, and let $b_\delta = 1_{\{v > \delta\}} b_E$. By Theorem 1.7 and by (1-16), (1-17), and (1-19), recalling also (2-10), (2-12) and the definition of $|D^c b_E|^+$, we have that $b_\delta \in GBV(\mathbb{R}^{n-1})$ with

$$\nabla b_\delta(z) = 0 \quad \text{for } \mathcal{H}^{n-1}\text{-a.e. } z \in \{v > \delta\}, \tag{3-95}$$

$$2[b_\delta] \le [v] \quad \mathcal{H}^{n-2}\text{-a.e. on } \{v > \delta\}^{(1)}, \tag{3-96}$$

$$2|D^c b_\delta|(G) \le |D^c v|(G) \quad \text{for every Borel set } G \subset \mathbb{R}^{n-1}. \tag{3-97}$$

Since $D^c v = 0$, we have that $|D^c b_\delta| = 0$ on Borel sets, by (3-97). Since, trivially, $\nabla b_\delta = 0$ $\mathcal{H}^{n-1}$-a.e. on $\{v \le \delta\}$, by (3-95) we have that $\nabla b_\delta = 0$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$. Finally, by (3-96) we have that

$$S_{b_\delta} \subset_{\mathcal{H}^{n-2}} (S_v \cap \{v > \delta\}^{(1)}) \cup \partial^e \{v > \delta\} \subset (S_v \cap \{v^\wedge > 0\}) \cup \partial^e \{v > \delta\}, \tag{3-98}$$

so that S_{b_δ} is locally $\mathcal{H}^{n-2}$-finite. We can thus apply Theorem 3.12 to b_δ to find a finite or countable set I_δ, $\{c_h^\delta\}_{h \in I_\delta} \subset \mathbb{R}$, and a Borel partition $\{G_h^\delta\}_{h \in I_\delta}$ of $\{v > \delta\}$ with

$$b_\delta = \sum_{h \in I_\delta} c_h^\delta 1_{G_h^\delta} \quad \mathcal{H}^{n-1}\text{-a.e. on } \{v > \delta\}.$$

By a diagonal argument over a sequence $\delta_h \to 0$ as $h \to \infty$ with $\{v > \delta_h\}$ of finite perimeter for every $h \in \mathbb{N}$, we prove the existence of I, $\{c_h\}_{h \in I}$ and $\{G_h\}_{h \in I}$ as in (1-22) such that $b_E = \sum_{h \in I} c_h 1_{G_h}$ $\mathcal{H}^{n-1}$-a.e. on $\{v > 0\}$ (and thus $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$). This means that

$$b_\delta = \sum_{h \in I_\delta} c_h 1_{G_h \cap \{v > \delta\}} \quad \mathcal{H}^{n-1}\text{-a.e. on } \mathbb{R}^{n-1},$$

and thus, again by Theorem 3.12, $\sum_{h \in I} P(G_h \cap \{v > \delta\} \cap B_R) < \infty$. This shows that $\{G_h\}_{h \in \mathbb{N}}$ is v-admissible and completes the proof.

Step two: We now assume that E is a v-distributed set of finite perimeter such that (1-22) holds, with $\{G_h\}_{h \in I}$ v-admissible, and $2[b_E] \le [v]$ $\mathcal{H}^{n-2}$-a.e. on $\{v^\wedge > 0\}$, and aim to prove that $E \in \mathcal{M}(v)$. Since E is v-distributed with segments as sections and $\{G_h\}_{h \in I}$ is v-admissible, we see that b_δ satisfies assumption (ii) of Theorem 3.12 for a.e. $\delta > 0$. By applying that theorem, and then by letting $\delta \to 0^+$, we deduce that $\nabla b_E = 0$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$ and that $|D^c b_E|^+ = 0$. Hence, by applying Theorem 1.9, we deduce that $E \in \mathcal{M}(v)$. $\square$

4. Rigidity in Steiner's inequality

In this section we discuss the rigidity problem for Steiner's inequality. We begin in Section 4A by proving the general sufficient condition for rigidity stated in Theorem 1.11. We then present our characterizations of rigidity: in Section 4B we prove Theorem 1.29 (characterization of rigidity for $v \in SBV(\mathbb{R}^{n-1}; [0, \infty))$ with S_v locally $\mathcal{H}^{n-2}$-finite), while Section 4C and 4E deal with the cases of generalized polyhedra and "no vertical boundaries". (Note that the equivalence between the indecomposability of $F[v]$ and the condition that $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$ is proved in Section 4D.) Finally, in Section 4F we address the proof of Theorem 1.30 about the characterization of equality cases for planar sets.

4A. *A general sufficient condition for rigidity.* The general sufficient condition of Theorem 1.11 follows quite easily from Theorem 1.9.

Proof of Theorem 1.11. Let $E \in \mathcal{M}(v)$, so that, by Theorem 1.9, we know that

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \partial^e\{b_E > t\})\, dt = \int_{G \cap S_{b_E} \cap S_v} [b_E]\, d\mathcal{H}^{n-2} + |D^c b_E|^+(G \cap K) \qquad (4\text{-}1)$$

whenever G is a Borel subset of $\{v^\wedge > 0\}$ and K is a Borel set of concentration for $|D^c b_E|^+$. If b_E is not constant on $\{v > 0\}$, then there exists a Lebesgue measurable set $I \subset \mathbb{R}$ such that $\mathcal{H}^1(I) > 0$ and, for every $t \in I$, the Borel sets $G_+ = \{b_E > t\} \cap \{v > 0\}$ and $G_- = \{b_E \le t\} \cap \{v > 0\}$ define a nontrivial Borel partition $\{G_+, G_-\}$ of $\{v > 0\}$. Since

$$\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- = \{v > 0\}^{(1)} \cap \partial^e\{b_E > t\},$$

by (1-21) we deduce that

$$\mathcal{H}^{n-2}\big(\big(\{v>0\}^{(1)}\cap\partial^e\{b_E>t\}\big)\setminus(\{v^\wedge=0\}\cup S_v\cup K)\big)>0 \quad\text{for all } t\in I. \qquad (4\text{-}2)$$

At the same time, by plugging $G=\{v>0\}^{(1)}\setminus(\{v^\wedge=0\}\cup S_v\cup K)\subset\{v^\wedge>0\}$ into (4-1), we find

$$\int_{\mathbb{R}}\mathcal{H}^{n-2}\big(\big(\{v>0\}^{(1)}\cap\partial^e\{b_E>t\}\big)\setminus(\{v^\wedge=0\}\cup S_v\cup K)\big)\,dt=0.$$

This is of course in contradiction with (4-2) and $\mathcal{H}^1(I)>0$. $\qquad\qquad\square$

Remark 4.1. By the same argument used in the proof of Theorem 1.11, one easily sees that if a Borel set $G\subset\mathbb{R}^m$ is essentially connected and $f\in BV(\mathbb{R}^m)$ is such that $|Df|(G^{(1)})=0$, then there exists $c\in\mathbb{R}$ such that $f=c$ $\mathcal{H}^m$-a.e. on G. In the case that G is an indecomposable set, this property was proved in [Dolzmann and Müller 1995, Proposition 2.12].

4B. *Characterization of rigidity for v in SBV with locally finite jump.* This section contains the proof of Theorem 1.29.

Proof of Theorem 1.29. Step one: We first prove that the mismatched stairway property implies rigidity. We argue by contradiction, and assume the existence of $E\in\mathcal{M}(v)$ such that $\mathcal{H}^n\big(E\triangle(te_n+F[v])\big)>0$ for every $t\in\mathbb{R}$. By Theorem 1.13, there exists a finite or countable set I, $\{c_h\}_{h\in I}\subset\mathbb{R}$, $\{G_h\}_{h\in I}$ a v-admissible partition of $\{v>0\}$ such that $b_E=\sum_{h\in I}c_h 1_{G_h}$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, $E=_{\mathcal{H}^n}W[v,b_E]$, and

$$2[b_E]\le[v]\quad\mathcal{H}^{n-2}\text{-a.e. on }\{v^\wedge>0\}. \qquad (4\text{-}3)$$

Of course, we may assume without loss of generality that $\mathcal{H}^{n-1}(G_h)>0$ for every $h\in I$ and that $c_h\ne c_k$ for every $h,k\in I$, $h\ne k$ (if any). In fact, $\#I\ge2$, because if $\#I=1$ then we would have $\mathcal{H}^n\big(E\triangle(ce_n+F[v])\big)=0$ for some $c\in\mathbb{R}$. We can apply the mismatched stairway property to I, $\{G_h\}_{h\in I}$ and $\{c_h\}_{h\in I}$, to find $h_0,k_0\in I$, $h_0\ne k_0$, and a Borel set Σ with $\mathcal{H}^{n-2}(\Sigma)>0$ such that

$$\Sigma\subset\partial^e G_{h_0}\cap\partial^e G_{k_0}\cap\{v^\wedge>0\}\qquad\text{and}\qquad[v](z)<2|c_{h_0}-c_{k_0}|\quad\text{for all } z\in\Sigma. \qquad (4\text{-}4)$$

Since $b_E^\vee\ge\max\{c_{h_0},c_{k_0}\}$ and $b_E^\wedge\le\min\{c_{h_0},c_{k_0}\}$ on $\partial^e G_{h_0}\cap\partial^e G_{k_0}$, (4-3) implies

$$2|c_{h_0}-c_{k_0}|\le[v]\quad\mathcal{H}^{n-2}\text{-a.e. on }\partial^e G_{h_0}\cap\partial^e G_{k_0}\cap\{v^\wedge>0\},$$

a contradiction to (4-4) and $\mathcal{H}^{n-2}(\Sigma)>0$.

Step two: We show that the failure of the mismatched stairway property implies the failure of rigidity. Indeed, let us assume the existence of a v-admissible partition $\{G_h\}_{h\in I}$ of $\{v>0\}$, and $\{c_h\}_{h\in I}\subset\mathbb{R}$ with $c_h\ne c_k$ for every $h,k\in I$, $h\ne k$, such that

$$2|c_h-c_k|\le[v]\quad\mathcal{H}^{n-2}\text{-a.e. on }\partial^e G_h\cap\partial^e G_k\cap\{v^\wedge>0\} \qquad (4\text{-}5)$$

whenever $h,k\in I$ with $h\ne k$. We now claim that $E\in\mathcal{M}(v)$, where

$$E=\bigcup_{h\in I}\big(c_h e_n+(F[v]\cap(G_h\times\mathbb{R}))\big).$$

To prove this claim, let $\delta > 0$ be such that $\{v > \delta\}$ is a set of finite perimeter. By Theorem 3.12, $b_\delta = b_E 1_{\{v > \delta\}} \in GBV(\mathbb{R}^{n-1})$ with $\nabla b_\delta = 0$ $\mathscr{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, $|D^c b_\delta| = 0$, S_{b_δ} is locally $\mathscr{H}^{n-2}$-finite, and

$$\{v > \delta\}^{(1)} \cap S_{b_\delta} \subset_{\mathscr{H}^{n-2}} \bigcup_{h \neq k \in I} \partial^e G_{h,\delta} \cap \partial^e G_{k,\delta}, \tag{4-6}$$

$$[b_\delta] = |c_h - c_k| \quad \mathscr{H}^{n-2}\text{-a.e. on } \partial^e G_{h,\delta} \cap \partial^e G_{k,\delta} \cap \{v > \delta\}^{(1)}, \, h \neq k \in I, \tag{4-7}$$

where $G_{h,\delta} = G_h \cap \{v > \delta\}$ for every $h \in I$. By (4-5), (4-6), and (4-7), we find

$$2[b_\delta] \leq [v] \quad \mathscr{H}^{n-2}\text{-a.e. on } S_{b_\delta} \cap \{v > \delta\}^{(1)}. \tag{4-8}$$

Now let $\{\delta_h\}_{h \in \mathbb{N}}$, $\{L_h\}_{h \in \mathbb{N}}$ be sequences satisfying (3-63), (3-64), and set $E_h = E \cap (\{\delta_h < v < L_h\} \times \mathbb{R})$, $\Sigma_h = \{\delta_h < v < L_h\}$, $b_h = 1_{\Sigma_h} b_E = 1_{\{v < L_h\}} b_{\delta_h}$ and $v_h = 1_{\Sigma_h} v$. Since $v_h \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ and $b_h \in GBV(\mathbb{R}^{n-1})$, we can apply Corollary 3.3 to compute $P(E_h; \Sigma_h^{(1)} \times \mathbb{R})$, to get (using that $\nabla b_\delta = 0$ $\mathscr{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, $|D^c b_\delta| = 0$, and (4-8)), that

$$P(E_h; \Sigma_h^{(1)} \times \mathbb{R}) = P(F[v]; \Sigma_h^{(1)} \times \mathbb{R}) \quad \text{for all } h \in \mathbb{N};$$

in particular,

$$\lim_{h \to \infty} P(E_h; \Sigma_h^{(1)} \times \mathbb{R}) = P(F[v]; \{v^\wedge > 0\} \times \mathbb{R}).$$

Moreover, by repeating the argument used in the proof of Proposition 3.8, we have

$$\lim_{h \to \infty} P(E_h; \partial^e \Sigma_h \times \mathbb{R}) = P(F[v]; \{v^\wedge = 0\} \times \mathbb{R}).$$

We thus conclude that

$$P(E) \leq \liminf_{h \to \infty} P(E_h) = P(F[v]),$$

that is, E is of finite perimeter with $E \in \mathcal{M}(v)$. $\qquad\square$

4C. *Characterization of rigidity on generalized polyhedra.* We now prove Theorem 1.20. The proof is based on the following lemma.

Lemma 4.2. *If $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathscr{H}^{n-1}(\{v > 0\}) < \infty$ is such that*

$$\{v > 0\} \text{ is of finite perimeter}, \tag{4-9}$$

$$\{v^\vee = 0\} \cap \{v > 0\}^{(1)} \text{ and } S_v \text{ are } \mathscr{H}^{n-2}\text{-finite}, \tag{4-10}$$

and if there exists $\varepsilon > 0$ such that $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$, then there exists $E \in \mathcal{M}(v)$ such that $\mathscr{H}^n\big(E \Delta (t e_n + F[v])\big) > 0$ for every $t \in \mathbb{R}$.

Proof. If $\varepsilon > 0$ is such that $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$, then there exists a nontrivial Borel partition $\{G_+, G_-\}$ of $\{v > 0\}$ modulo $\mathscr{H}^{n-1}$ such that

$$\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathscr{H}^{n-2}} \{v^\wedge = 0\} \cup \{[v] > \varepsilon\}. \tag{4-11}$$

We are now going to show that the set E defined by

$$E = \big((\tfrac{1}{2}\varepsilon e_n + F[v]) \cap (G_+ \times \mathbb{R})\big) \cup (F[v] \cap (G_- \times \mathbb{R}))$$

satisfies $E \in \mathcal{M}(v)$; this will prove the lemma. To this end we first prove that G_+ is a set of finite perimeter. Indeed, since $G_+ \subset \{v > 0\}$, we have

$$\partial^e G_+ \subset (\partial^e G_+ \cap \{v > 0\}^{(1)}) \cup \partial^e \{v > 0\}, \tag{4-12}$$

where $\partial^e G_+ \cap \{v > 0\}^{(1)} = \partial^e G_+ \cap \partial^e G_- \cap \{v > 0\}^{(1)}$, and thus, by (4-11),

$$\partial^e G_+ \cap \{v > 0\}^{(1)} \subset_{\mathcal{H}^{n-2}} \partial^e G_+ \cap \{v > 0\}^{(1)} \cap (\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}) \subset (\partial^e G_+ \cap \{v^\vee = 0\} \cap \{v > 0\}^{(1)}) \cup S_v. \tag{4-13}$$

By combining (4-9), (4-10) (4-12), and (4-13), we conclude that $\mathcal{H}^{n-2}(\partial^e G_+) < \infty$, and thus, by Federer's criterion, that G_+ is a set of finite perimeter. Since $b_E = \frac{1}{2}\varepsilon 1_{G_+}$, we thus have $b_E \in BV(\mathbb{R}^{n-1})$, and thus $E = W[v, b_E]$ is of finite perimeter with segments as sections. Since $\nabla b_E = 0$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$ and $D^c b_E = 0$, we are only left to check that $2[b_E] \leq [v]$ $\mathcal{H}^{n-2}$-a.e. on $\{v^\wedge > 0\}$ in order to conclude that $E \in \mathcal{M}(v)$ by means of Theorem 1.9. Indeed, since $b_E = \frac{1}{2}\varepsilon 1_{G_+}$, we have $S_{b_E} = \partial^e G_+$ with $[b_E] = \frac{1}{2}\varepsilon$ $\mathcal{H}^{n-2}$-a.e. on $\partial^e G_+$. By (2-9) and (4-11),

$$S_{b_E} \cap \{v^\wedge > 0\} = \partial^e G_+ \cap \{v^\wedge > 0\} = \partial^e G_+ \cap \partial^e G_- \cap \{v > 0\}^{(1)} \cap \{v^\wedge > 0\} \subset_{\mathcal{H}^{n-2}} \{[v] > \varepsilon\}. \qquad \square$$

Proof of Theorem 1.20. Step one: We prove that, if $F[v]$ is a generalized polyhedron, then $v \in SBV(\mathbb{R}^{n-1})$, S_v and $\{v^\vee = 0\} \setminus \{v = 0\}^{(1)}$ are $\mathcal{H}^{n-2}$-finite, and $\{v > 0\}$ is of finite perimeter. Indeed, by assumption, there exist a finite disjoint family of indecomposable sets of finite perimeter and volume $\{A_j\}_{j \in J}$ in $\mathbb{R}^{n-1}$, and a family of functions $\{v_j\}_{j \in J} \subset W^{1,1}(\mathbb{R}^{n-1})$, such that

$$v = \sum_{j \in J} v_j 1_{A_j}, \quad (\{v^\wedge = 0\} \setminus \{v = 0\}^{(1)}) \cup S_v \subset_{\mathcal{H}^{n-2}} \bigcup_{j \in J} \partial^e A_j. \tag{4-14}$$

By [Ambrosio et al. 2000, Example 4.5], $v_j 1_{A_j} \in SBV(\mathbb{R}^{n-1})$ for every $j \in J$, so that $v \in SBV(\mathbb{R}^{n-1})$, as J is finite. Similarly, (4-14) gives that $\{v^\wedge = 0\} \setminus \{v = 0\}^{(1)}$ and S_v are both $\mathcal{H}^{n-2}$-finite. Since $\{v^\vee = 0\} \setminus \{v = 0\}^{(1)}$ and $\partial^e \{v > 0\}$ are both subsets of $\{v^\wedge = 0\} \setminus \{v = 0\}^{(1)}$, we deduce that $\{v^\vee = 0\} \setminus \{v = 0\}^{(1)}$ and $\partial^e \{v > 0\}$ are $\mathcal{H}^{n-2}$-finite. In particular, by Federer's criterion, $\{v > 0\}$ is a set of finite perimeter.

Step two: By step one, if $F[v]$ is a generalized polyhedron, then v satisfies the assumptions of Lemma 4.2. In particular, if $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$, then rigidity fails. This shows the implication (i) $\Rightarrow$ (ii) in the theorem.

Step three: We show that if rigidity fails, then $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$. By step one, if $F[v]$ is a generalized polyhedron, then v satisfies the assumptions of Theorem 1.13. In particular, if $E \in \mathcal{M}(v)$, then $\nabla b_E = 0$, $S_{b_E} \cap \{v^\wedge > 0\} \subset S_v$, $2[b_E] \leq [v]$ $\mathcal{H}^{n-2}$-a.e. on $\{v^\wedge > 0\}$, and $|D^c b_E|^+ = 0$, so that, by (1-28) and (1-20), we find

$$S_{b_E} \subset_{\mathcal{H}^{n-2}} \bigcup_{j \in J} \partial^e A_j, \tag{4-15}$$

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(G \cap \partial^e \{b_E > t\}) \, dt = \int_{G \cap S_{b_E}} [b_E] \, d\mathcal{H}^{n-2} \tag{4-16}$$

for every Borel set $G \subset \{v^\wedge > 0\}$. We now combine (4-15) and (4-16) to deduce that

$$\int_{\mathbb{R}} \mathcal{H}^{n-2}(A_j^{(1)} \cap \partial^e \{b_E > t\}) \, dt = 0 \quad \text{for all } j \in J.$$

Since each A_j is indecomposable, by arguing as in the proof of Theorem 1.11 we see that there exists $\{c_j\}_{j \in J} \subset \mathbb{R}$ such that $b_E = \sum_{j \in J} c_j 1_{A_j}$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$. In particular, we have $b_E = \sum_{j \in J_0} a_j 1_{B_j}$ $\mathcal{H}^{n-1}$-a.e. on $\mathbb{R}^{n-1}$, where $\#J_0 \leq \#J$, $\{a_j\}_{j \in J_0} \subset \mathbb{R}$ with $a_j \neq a_i$ if $i, j \in J_0$, $i \neq j$, and $\{B_j\}_{j \in J_0}$ is a partition modulo $\mathcal{H}^{n-1}$ of $\mathbb{R}^{n-1}$ into sets of finite perimeter. (Notice that each B_j may fail to be indecomposable.) Let us now assume, in addition to $E \in \mathcal{M}(v)$, that $\mathcal{H}^n(E \Delta (te_n + F[v])) > 0$ for every $t \in \mathbb{R}$. In this case, the formula for b_E we have just proved implies that $\#J_0 \geq 2$. We now set

$$\varepsilon = \min\{|a_i - a_j| : i, j \in J_0, \ i \neq j\},$$

so that $\varepsilon > 0$, and, for some $j_0 \in J_0$, we set $G_+ = B_{j_0}$ and $G_- = \bigcup_{j \in J_0, j \neq j_0} B_j$. In this way $\{G_+, G_-\}$ defines a nontrivial Borel partition of $\{v > 0\}$ modulo $\mathcal{H}^{n-1}$ with the property that

$$[v] \geq 2[b_E] \geq 2\varepsilon \quad \mathcal{H}^{n-2}\text{-a.e. on } \{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_- .$$

Thus, $\{v^\wedge = 0\} \cup \{[v] > \varepsilon\}$ essentially disconnects $\{v > 0\}$, and the proof of Theorem 1.20 is complete. $\square$

4D. *Characterization of indecomposability on Steiner symmetrals.* We show here that requiring that $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$ is in fact equivalent to saying that $F[v]$ is an indecomposable set of finite perimeter. This result shall be used to provide a second type of characterization of rigidity when $F[v]$ has no vertical parts, as well as in the planar case; see Theorem 1.16 and Theorem 1.30.

Theorem 4.3. *If $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, then $F[v]$ is indecomposable if and only if $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$.*

We start by recalling a version of Vol'pert's theorem; see [Barchiesi et al. 2013, Theorem 2.4].

Theorem C. *If E is a set of finite perimeter in $\mathbb{R}^n$, then there exists a Borel set $G_E \subset \{v > 0\}$ with $\mathcal{H}^{n-1}(\{v > 0\} \setminus G_E) = 0$ such that E_z is a set of finite perimeter in $\mathbb{R}$ with $\partial^*(E_z) = (\partial^* E)_z$ for every $z \in G_E$. Moreover, if $z \in G_E$ and $s \in \partial^* E_z$, then*

$$\boldsymbol{q}\nu_E(z, s) \neq 0, \quad \nu_{E_z}(s) = \frac{\boldsymbol{q}\nu_E(z, s)}{|\boldsymbol{q}\nu_E(z, s)|} . \tag{4-17}$$

Proof of Theorem 4.3. In Lemma 4.4 below, we prove that, if $F = F[v]$ is indecomposable, then $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$. We prove here the reverse implication. Precisely, let us assume the existence of a nontrivial partition $\{F_+, F_-\}$ of F into sets of finite perimeter such that

$$0 = \mathcal{H}^{n-1}(F^{(1)} \cap \partial^e F_+ \cap \partial^e F_-) = \mathcal{H}^{n-1}(F^{(1)} \cap \partial^e F_+). \tag{4-18}$$

We aim to prove that, if we set

$$G_+ = \{z \in \mathbb{R}^{n-1} : \mathcal{H}^1((F_+)_z) > 0\}, \quad G_- = \{z \in \mathbb{R}^{n-1} : \mathcal{H}^1((F_-)_z) > 0\},$$

then $\{G_+, G_-\}$ defines a nontrivial Borel partition modulo $\mathcal{H}^{n-1}$ of $\{v > 0\}$ such that

$$\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\}. \tag{4-19}$$

Step one: We prove that $\{G_+, G_-\}$ is a nontrivial Borel partition (modulo $\mathcal{H}^{n-1}$) of $\{v > 0\}$. The only nontrivial fact to obtain is that $\mathcal{H}^{n-1}(G_+ \cap G_-) = 0$. By Theorem C there exists $G_+^* \subset G_+$ with $\mathcal{H}^{n-1}(G_+ \setminus G_+^*) = 0$ such that, if $z \in G_+^*$, then

- $(F_+)_z$ is a set of finite perimeter in $\mathbb{R}$ with $(\partial^* F_+)_z = \partial^*((F_+)_z)$,
- $(F_-)_z$ is a set of finite perimeter in $\mathbb{R}$,
- $\{(F_+)_z, (F_-)_z\}$ is a partition modulo $\mathcal{H}^1$ of $(F^{(1)})_z$,

where the last property follows by Fubini's theorem and $\mathcal{H}^n(F \Delta F^{(1)}) = 0$. Now let

$$G_+^{**} = \{z \in G_+^* : \mathcal{H}^1((F^{(1)})_z \setminus (F_+)_z) > 0\} = G_+^* \cap G_-\,.$$

If $z \in G_+^{**}$, then $\{(F_+)_z, (F_-)_z\}$ is a nontrivial partition modulo $\mathcal{H}^1$ of $(F^{(1)})_z$ into sets of finite perimeter. Since $(F^{(1)})_z$ is an interval for every $z \in \mathbb{R}^{n-1}$ (see [Maggi 2012, Lemma 14.6]), we thus have

$$\mathcal{H}^0\big([(F^{(1)})_z]^{(1)} \cap \partial^*((F_+)_z) \cap \partial^*((F_-)_z)\big) \geq 1 \quad \text{for all } z \in G_+^{**}.$$

In particular, since $(\partial^* F_+)_z = \partial^*((F_+)_z)$, $[(F^{(1)})_z]^{(1)} \subset (F^{(1)})_z$, and $(A \cap B)_z = A_z \cap B_z$ for every $A, B \subset \mathbb{R}^n$, we have

$$\mathcal{H}^0((F^{(1)} \cap \partial^* F_+)_z) \geq 1 \quad \text{for all } z \in G_+^{**}.$$

Hence, $G_+^{**} \subset p(F^{(1)} \cap \partial^* F_+)$, and by (4-18) and [Maggi 2012, Proposition 3.5] we conclude

$$0 = \mathcal{H}^{n-1}(F^{(1)} \cap \partial^* F_+) \geq \mathcal{H}^{n-1}(p(F^{(1)} \cap \partial^* F_+)) \geq \mathcal{H}^{n-1}(G_+^{**}) = \mathcal{H}^{n-1}(G_+^* \cap G_-),$$

that is, $\mathcal{H}^{n-1}(G_+ \cap G_-) = 0$.

Step two: We now show that

$$F^{(1)} \cap ((\partial^e G_+ \cap \partial^e G_-) \times \mathbb{R}) \subset \partial^e F_+ \cap \partial^e F_-\,. \tag{4-20}$$

Indeed, let (z, s) belong to the set on the left-hand side of this inclusion; if — seeking contradiction — $(z, s) \notin \partial^e F_+ \cap \partial^e F_-$, then either $(z, s) \in F_-^{(1)}$ or $(z, s) \in F_+^{(1)}$. In the former case,

$$\mathcal{H}^n(C_{(z,s),r}) = \mathcal{H}^n(F_- \cap C_{(z,s),r}) + o(r^n) \leq 2r \mathcal{H}^{n-1}(G_- \cap D_{z,r}) + o(r^n),$$

that is, $z \in G_-^{(1)}$, contradicting $z \in \partial^e G_-$; the latter case is treated analogously.

Step three: We conclude the proof. Arguing by contradiction, we can assume that

$$0 < \mathcal{H}^{n-2}(\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \setminus \{v^\wedge = 0\})$$
$$= \mathcal{H}^{n-2}(\partial^e G_+ \cap \partial^e G_- \cap \{v^\wedge > 0\})$$
$$= \lim_{\varepsilon \to 0^+} \mathcal{H}^{n-2}(\partial^e G_+ \cap \partial^e G_- \cap \{v^\wedge > \varepsilon\}),$$

where it should be noted that all these measures could be equal to $+\infty$. However, by [Mattila 1995, Theorem 8.13], if ε is sufficiently small, then there exists a compact set K with $0 < \mathcal{H}^{n-2}(K) < \infty$ and $K \subset \partial^e G_+ \cap \partial^e G_- \cap \{v^\wedge > \varepsilon\}$. Therefore, by (4-20),

$$
\begin{aligned}
\mathcal{H}^{n-1}(F^{(1)} \cap \partial^e F_+ \cap \partial^e F_-) &\geq \mathcal{H}^{n-1}\big(F^{(1)} \cap ((\partial^e G_+ \cap \partial^e G_-) \times \mathbb{R})\big) \\
&\geq \mathcal{H}^{n-1}(F^{(1)} \cap (K \times \mathbb{R})) \\
&\geq \mathcal{H}^{n-1}\big(\{x \in \mathbb{R}^n : \boldsymbol{p}x \in K, |\boldsymbol{q}x| < \tfrac{1}{2}v^\wedge(\boldsymbol{p}x)\}\big) \quad \text{by (3-27)} \\
&\geq \mathcal{H}^{n-1}(\{x \in \mathbb{R}^n : \boldsymbol{p}x \in K, |\boldsymbol{q}x| < \tfrac{1}{2}\varepsilon\}) \qquad \text{since } K \subset \{v^\wedge > \varepsilon\} \\
&\geq c(n)\mathcal{H}^{n-2}(K)\varepsilon > 0
\end{aligned}
$$

by [Federer 1969, 2.10.45], a contradiction to (4-18). $\square$

Lemma 4.4. *Let* $v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ *with* $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$. *If* $\{G_+, G_-\}$ *is a Borel partition of* $\{v > 0\}$ *such that*

$$\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\}, \tag{4-21}$$

then $F_+ = F[v] \cap (G_+ \times \mathbb{R})$ *and* $F_- = F[v] \cap (G_- \times \mathbb{R})$ *are sets of finite perimeter, with*

$$P(F_+) + P(F_-) = P(F[v]).$$

Proof. Step one: We prove that F_+ is a set of finite perimeter (the same argument works, of course, in the case of F_-). Indeed, let $G_{+0} = G_+ \cup \{v = 0\}$. Since $F[v] \cap (G_{+0} \times \mathbb{R}) = F_+ \cap (G_{+0} \times \mathbb{R})$, we find that

$$\mathcal{H}^{n-1}(\partial^e F \cap (G_{+0}^{(1)} \times \mathbb{R})) = \mathcal{H}^{n-1}(\partial^e F_+ \cap (G_{+0}^{(1)} \times \mathbb{R})), \tag{4-22}$$

where we have set $F = F[v]$. Since $\partial^e F_+ \cap (G_{+0}^{(0)} \times \mathbb{R}) = \varnothing$, we find

$$\mathcal{H}^{n-1}(\partial^e F_+ \cap (G_{+0}^{(0)} \times \mathbb{R})) = 0. \tag{4-23}$$

We now note that

$$\mathbb{R}^{n-1} \setminus (G_{+0}^{(1)} \cup G_{+0}^{(0)}) = \partial^e G_{+0} = \partial^e G_- .$$

Since $\{v > 0\}^{(0)} \cap \partial^e G_- = \varnothing$, $\partial^e\{v > 0\} \subset \{v^\wedge = 0\}$, and $\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- = \{v > 0\}^{(1)} \cap \partial^e G_-$, by (4-21) we find that

$$\partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\}. \tag{4-24}$$

Thus, by (4-22), (4-23), (4-24), and by Federer's criterion, in order to prove that F_+ is a set of finite perimeter, we are left to show that

$$\mathcal{H}^{n-1}\big(\partial^e F_+ \cap (\{v^\wedge = 0\} \times \mathbb{R})\big) < \infty. \tag{4-25}$$

Since $(\partial^e F_+)_z = \varnothing$ whenever $z \in \{v = 0\}^{(1)}$, we find that

$$\mathcal{H}^{n-1}\big(\partial^e F_+ \cap (\{v = 0\}^{(1)} \times \mathbb{R})\big) = 0. \tag{4-26}$$

Since $F_+ \subset F$, $\partial^e F_+ \subset F^{(1)} \cup \partial^e F$. At the same time, if $z \in \{v^\vee = 0\}$, then $(\partial^e F)_z \cup (F^{(1)})_z \subset \{0\}$ by

(3-27) and (3-28), so that, if $G \subset \{v^\vee = 0\}$, then

$$\mathcal{H}^{n-1}(\partial^e F_+ \cap (G \times \mathbb{R})) \leq \mathcal{H}^{n-1}(G \times \{0\}) = \mathcal{H}^{n-1}(G).$$

By the Lebesgue density theorem, $\mathcal{H}^{n-1}(\{v^\vee = 0\} \setminus \{v = 0\}^{(1)}) = 0$, thus, if we plug in the above identity $G = \{v^\vee = 0\} \setminus \{v = 0\}^{(1)}$, then (4-26) gives

$$\mathcal{H}^{n-1}\big(\partial^e F_+ \cap (\{v^\vee = 0\} \times \mathbb{R})\big) = 0. \tag{4-27}$$

Finally, if $z \in \{v^\wedge = 0 < v^\vee\}$, then $(F^{(1)})_z \subset \{0\}$ and $(\partial^e F)_z \subset [-\frac{1}{2}v^\vee(z), \frac{1}{2}v^\vee(z)]$ by Corollary 3.4. Since $\{v^\wedge = 0 < v^\vee\}$ is countably $\mathcal{H}^{n-2}$-rectifiable, by [Federer 1969, 3.2.23] and (3-29) we find

$$\mathcal{H}^{n-1}(\partial^e F_+ \cap (G \times \mathbb{R})) = \int_G \mathcal{H}^1((\partial^e F_+)_z) \, d\mathcal{H}^{n-2}(z) \leq \int_G v^\vee \, d\mathcal{H}^{n-2} = P(F; G \times \mathbb{R}) \tag{4-28}$$

for every Borel set $G \subset \{v^\wedge = 0 < v^\vee\}$. By combining (4-28) (with $G = \{v^\wedge = 0 < v^\vee\}$) and (4-27), we obtain (4-25) for F_+. The proof for F_- is of course entirely analogous.

Step two: We now prove that $P(F_+) + P(F_-) = P(F)$. Since F is $\mathcal{H}^n$-equivalent to $F_+ \cup F_-$, by [Maggi 2012, Lemma 12.22] it suffices to prove that $P(F_+) + P(F_-) \leq P(F)$. By (4-22), (4-27), and the analogous relations for F_-, we are actually left to show that

$$P(F_+; G \times \mathbb{R}) + P(F_-; G \times \mathbb{R}) \leq P(F; G \times \mathbb{R}) \tag{4-29}$$

for every Borel set $G \subset \{v^\wedge = 0 < v^\vee\}$. Since $F_+ = F[1_{G_+} v]$ is of finite perimeter, by Corollary 3.4 we have $v_+ = 1_{G_+} v \in BV(\mathbb{R}^{n-1})$, with

$$P(F_+; G \times \mathbb{R}) = 2 \int_{G \cap \{v_+ > 0\}} \sqrt{1 + |\tfrac{1}{2}\nabla v_+|^2} + \int_{G \cap S_{v_+}} [v_+] \, d\mathcal{H}^{n-2} + |D^c v_+|(G) \tag{4-30}$$

for every Borel set $G \subset \mathbb{R}^{n-1}$. Since $\{v^\wedge = 0 < v^\vee\}$ is countably $\mathcal{H}^{n-2}$-rectifiable, we find

$$P(F_+; G \times \mathbb{R}) = \int_{G \cap S_{v_+}} [v_+] \, d\mathcal{H}^{n-2} = P(F_+; G \cap S_{v_+})$$

for every Borel set $G \subset \{v^\wedge = 0 < v^\vee\}$; moreover, an analogous formula holds for F_-. Thus, (4-29) takes the form

$$P(F_+; G \cap S_{v_+}) + P(F_-; G \cap S_{v_-}) \leq P(F; G \times \mathbb{R}) \tag{4-31}$$

for every Borel set $G \subset \{v^\wedge = 0 < v^\vee\}$. If $G \subset \{v^\wedge = 0 < v^\vee\} \setminus S_{v_-}$, then (4-31) reduces to $P(F_+; G \cap S_{v_+}) \leq P(F; G \times \mathbb{R})$, which follows immediately from (4-28). A similar argument holds if we choose $G \subset \{v^\wedge = 0 < v^\vee\} \setminus S_{v_+}$. We may thus conclude the proof of the lemma by showing that

$$\mathcal{H}^{n-2}(\{v^\wedge = 0 < v^\vee\} \cap S_{v_+} \cap S_{v_-}) = 0. \tag{4-32}$$

To prove (4-32), let us note that for $\mathcal{H}^{n-2}$-a.e. $z \in \{v^\wedge = 0 < v^\vee\} \cap S_{v_+} \cap S_{v_-}$, we have

$$\{v > t\}_{z,t} \xrightarrow{\text{loc}} H_0 \quad \text{for all } t \in (0, v^\vee(z)), \tag{4-33}$$

$$\{v_+ > t\}_{z,t} \xrightarrow{\text{loc}} H_1 \quad \text{for all } t \in (v_+^\wedge(z), v_+^\vee(z)), \tag{4-34}$$

$$\{v_- > t\}_{z,t} \xrightarrow{\text{loc}} H_2 \quad \text{for all } t \in (v_-^\wedge(z), v_-^\vee(z))$$

as $r \to 0^+$. Now, $v_+^\vee(z) \le v^\vee(z)$, therefore $(v_+^\wedge(z), v_+^\vee(z)) \subset (0, v^\vee(z))$. We may thus pick $t > 0$ such that (4-33) and (4-34) hold, and, therefore,

$$\{v > t\}_{z,t} \xrightarrow{\text{loc}} H_0, \quad (G_+ \cap \{v > t\})_{z,r} = \{v_+ > t\}_{z,t} \xrightarrow{\text{loc}} H_1$$

as $r \to 0^+$. Since $G_+ \cap \{v > t\} \subset \{v > t\}$, we have $H_1 \subset H_0$, and thus $H_1 = H_0$. This implies that

$$\mathcal{H}^{n-1}\big(\boldsymbol{D}_{z,r} \cap ((z + H_0) \setminus G_+)\big) = o(r^{n-1}) \quad \text{as } r \to 0^+.$$

The same argument applies to v_- and gives

$$\mathcal{H}^{n-1}\big(\boldsymbol{D}_{z,r} \cap ((z + H_0) \setminus G_-)\big) = o(r^{n-1}) \quad \text{as } r \to 0^+.$$

Hence, $\theta^*(G_+ \cap G_-, z) \ge \theta(z + H_0, z) = \frac{1}{2}$, a contradiction to $\mathcal{H}^{n-1}(G_+ \cap G_-) = 0$. $\qquad \square$

4E. *Characterizations of rigidity without vertical boundaries.* We now prove Theorem 1.16, by combining Theorem 1.11 and the results from Section 4D.

Proof of Theorem 1.16. We start by noticing that the equivalence between (ii) and (iii) was proved in Theorem 4.3. We are thus left to prove the equivalence between (i) and (ii).

Step one: We prove that (ii) implies (i). By Lemma 2.2, we have that $D^c v \llcorner \{v^\wedge = 0\} = 0$; since we are now assuming that $D^s v \llcorner \{v^\wedge > 0\} = 0$, we conclude that $D^c v = 0$. We now show that $\{v^\wedge = 0\} \cup S_v$ does not essentially disconnect $\{v > 0\}$. Otherwise, there exists a nontrivial Borel partition $\{G_+, G_-\}$ modulo $\mathcal{H}^{n-1}$ of $\{v > 0\}$ such that

$$\{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_- \subset \{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\} \cup S_v, \tag{4-35}$$

where the first inclusion follows from (2-9). Since $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$ and since $D^s v \llcorner \{v^\wedge > 0\} = 0$ implies $\mathcal{H}^{n-2}(S_v \cap \{v^\wedge > 0\}) = 0$, we conclude

$$0 < \mathcal{H}^{n-2}\big((\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_-) \setminus \{v^\wedge = 0\}\big)$$
$$= \mathcal{H}^{n-2}(\{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_-) = \mathcal{H}^{n-2}\big((\{v^\wedge > 0\} \cap \partial^e G_+ \cap \partial^e G_-) \setminus S_v\big),$$

a contradiction to (4-35). This proves that $\{v^\wedge = 0\} \cup S_v$ does not essentially disconnect $\{v > 0\}$. Since $D^c v = 0$, we can thus apply Theorem 1.11 to deduce (i).

Step two: We prove that (i) implies (ii). Indeed, if (ii) fails, then there exists a nontrivial Borel partition $\{G_+, G_-\}$ of $\{v > 0\}$ modulo $\mathcal{H}^{n-1}$ such that $\{v > 0\}^{(1)} \cap \partial^e G_+ \cap \partial^e G_- \subset_{\mathcal{H}^{n-2}} \{v^\wedge = 0\}$. By Lemma 4.4, we find that $F_+ = F \cap (G_+ \times \mathbb{R})$ and $F_- = F \cap (G_- \times \mathbb{R})$ are sets of finite perimeter with

$P(F_+) + P(F_-) = P(F)$. Let us now set $E = (e_n + F_+) \cup F_-$. By [Maggi 2012, Lemma 12.22], we have that E is a v-distributed set of finite perimeter with

$$P(F) \le P(E) \le P(e_n + F_+) + P(F_-) = P(F_+) + P(F_-) = P(F),$$

that is, $E \in \mathcal{M}(v)$. However, $\mathcal{H}^n(E \Delta (te_n + F)) > 0$ for every $t \in \mathbb{R}$, since $\{G_+, G_-\}$ was a nontrivial Borel partition of $\{v > 0\}$. $\square$

4F. *Characterizations of rigidity on planar sets.* We finally prove Theorem 1.30, which addresses the rigidity problem for planar sets.

Proof of Theorem 1.30. Step one: Let us assume that (ii) holds. We first note that, in this case, $D^c v = 0$, so that, thanks to Theorem 1.11, we are left to prove that

$$\{v^\wedge = 0\} \cup S_v \text{ does not essentially disconnect } \{v > 0\} \tag{4-36}$$

in order to show the validity of (i). Since (ii) implies that $\{v^\wedge = 0\} \cup S_v \subset \mathbb{R} \setminus (a, b)$, where $\{v > 0\}$ is $\mathcal{H}^1$-equivalent to (a, b), (4-36) follows from the fact that $\mathbb{R} \setminus (a, b)$ does not essentially disconnect (a, b).

Step two: We now assume the validity of (i). Let $[a, b]$ be the least closed interval which contains $\{v > 0\}$ modulo $\mathcal{H}^1$. (Note that $[a, b]$ could a priori be unbounded.) Let us assume without loss of generality that $\mathcal{H}^1(\{v > 0\}) > 0$, so that (a, b) is nonempty. We now show that $v^\wedge(c) > 0$ for every $c \in (a, b)$. Indeed, let $F = F[v]$, $F_+ = F \cap [[c, \infty) \times \mathbb{R}]$, and $F_- = F \cap [(-\infty, c) \times \mathbb{R}]$. Since $F_+ = F[1_{[c,\infty)}v]$ and $F_- = F[1_{(-\infty,c)}v]$, we can apply (3-29) to find that

$$P(F_+) = 2 \int_{\{v>0\} \cap (c,\infty)} \sqrt{1 + |\tfrac{1}{2}v'|^2} + \int_{S_v \cap (c,\infty)} [v] \, d\mathcal{H}^0 + v(c^+) + |D^c v|(\{\tilde{v} > 0\} \cap (c, \infty)) \tag{4-37}$$

and

$$P(F_-) = 2 \int_{\{v>0\} \cap (-\infty,c)} \sqrt{1 + |\tfrac{1}{2}v'|^2} + \int_{S_v \cap (-\infty,c)} [v] \, d\mathcal{H}^0 + v(c^-) + |D^c v|(\{\tilde{v} > 0\} \cap (-\infty, c)), \tag{4-38}$$

where we have set $v(c^+) = \text{aplim}(v, (c, \infty), c)$, $v(c^-) = \text{aplim}(v, (-\infty, c), c)$, and we have used the fact that $D^c(1_{(c,\infty)}v)$ is the restriction of $D^c v$ to (c, ∞), that

$$[1_{(c,\infty)}v](z) = \begin{cases} [v](z) & \text{if } z > c, \\ v(c^+) & \text{if } z = c, \\ 0 & \text{if } z < c, \end{cases}$$

as well as the analogous facts for $1_{(-\infty,c)}v$. Notice that, if $v^\wedge(c) = 0$, then either $v(c^+) = 0$ or $v(c^-) = 0$, and, therefore, $P(F_+) + P(F_-) = P(F)$ by (3-29), (4-37), and (4-38). As a consequence, if we set $E = F_+ \cup (e_2 + F_-)$, then by arguing as in step two of the proof of Theorem 1.16 we find that

$$P(F) \le P(E) \le P(F_+) + P(e_2 + F_-) = P(F_+) + P(F_-) = P(F),$$

that is, $E \in \mathcal{M}(v)$, in contradiction to (i). This proves that $v^\wedge(c) > 0$ for every $c \in (a, b)$. In particular, since $\{v > 0\}$ is $\mathcal{H}^1$-equivalent to $\{v^\wedge > 0\}$, we find that $\{v > 0\}$ is $\mathcal{H}^1$-equivalent to (a, b). We now prove that (a, b) is bounded. Let us decompose v as $v = v_1 + v_2$, where $v_1 \in W^{1,1}(\mathbb{R})$ and $v_2 \in BV(\mathbb{R})$ with

$D^a v_2 = 0$; see [Ambrosio et al. 2000, Corollary 3.33]. If v_2 is nonconstant (modulo $\mathcal{H}^1$) in (a, b), then we find a contradiction with (i), by Proposition 1.15. Thus, there exists $t \in \mathbb{R}$ such that $v_2 = t$ on (a, b), and so $v = v_1 + t \in W^{1,1}(a, b)$. In particular, since $\{v > 0\} =_{\mathcal{H}^1} (a, b)$ and $\mathcal{H}^1(\{v > 0\}) < \infty$, we find that (a, b) is bounded.

Step three: We prove that (ii) implies (iii). Indeed, since $\{v > 0\}$ is $\mathcal{H}^1$-equivalent to (a, b) and $v^\wedge > 0$ on (a, b), by Remark 1.5 we have that $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$. In particular, by Theorem 4.3, we have that $F[v]$ is indecomposable. Since $v \in W^{1,1}(a, b)$, by [Chlebík et al. 2005, Proposition 1.2], we find that

$$\mathcal{H}^1\big(\{x \in \partial^* F[v] : \boldsymbol{q} v_{F[v]} = 0, \, \boldsymbol{p} x \in (a, b)\}\big) = 0. \tag{4-39}$$

Since $\{v^\wedge > 0\} = (a, b)$, we deduce (1-33).

Step four: We prove that (iii) implies (ii). Since $F[v]$ is now indecomposable, by Theorem 4.3 we have that $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$. In particular, $\{v > 0\}$ is an essentially connected subset of $\mathbb{R}$, and thus, by [Cagnetti et al. 2013, Proof of Theorem 1.6, step one], $\{v > 0\}$ is $\mathcal{H}^1$-equivalent to an interval. Since $\mathcal{H}^1(\{v > 0\}) < \infty$, we thus have that $\{v > 0\} =_{\mathcal{H}^1} (a, b)$, with (a, b) bounded. Since $\{v^\wedge = 0\}$ does not essentially disconnect $\{v > 0\}$, we have $v^\wedge > 0$ on (a, b). Finally, by (1-33) and the fact that $v^\wedge > 0$ on (a, b), we find (4-39). Again by [Chlebík et al. 2005, Proposition 1.2], we conclude that $v \in W^{1,1}(a, b)$. $\qquad\square$

Appendix A: Equality cases in the localized Steiner inequality

The rigidity results described in this paper for the equality cases in Steiner's inequality $P(E) \ge P(F[v])$ can be suitably formulated and proved for the localized Steiner inequality $P(E; \Omega \times \mathbb{R}) \ge P(F[v]; \Omega \times \mathbb{R})$ under the assumption that Ω is an open connected set. This generalization does not require the introduction of new ideas, but, of course, requires clumsier notation. Another possible approach is that of obtaining the localized rigidity results through an approximation process. For the sake of clarity, we exemplify this by showing a proof of Theorem B based on Theorem 1.11. The required approximation technique is described in the following lemma.

Lemma A.1. *If Ω is a connected open set in $\mathbb{R}^{n-1}$, $v \in BV(\Omega; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, E is a v-distributed set with $P(E; \Omega \times \mathbb{R}) < \infty$ and segments as vertical sections, then there exists an increasing sequence $\{\Omega_k\}_{k \in \mathbb{N}}$ of bounded open connected sets of finite perimeter such that $\Omega = \bigcup_{k \in \mathbb{N}} \Omega_k$, Ω_k is compactly contained in Ω, $v_k = 1_{\Omega_k} v \in BV(\mathbb{R}^{n-1}; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v_k > 0\}) < \infty$, $E_k = E \cap (\Omega_k \times \mathbb{R})$ is a v_k-distributed set of finite perimeter, and*

$$P(E_k) = P(E; \Omega_k \times \mathbb{R}) + P(F[v_k]; \partial^* \Omega_k \times \mathbb{R}), \tag{A-1}$$

$$P(F[v_k]) = P(F[v]; \Omega_k \times \mathbb{R}) + P(F[v_k]; \partial^* \Omega_k \times \mathbb{R}). \tag{A-2}$$

Finally, if $E \in \mathcal{M}_\Omega(v)$ — see (1-2) — then $E_k \in \mathcal{M}(v_k)$.

Proof. By intersecting Ω with increasingly larger balls, and by a diagonal argument, we may assume that Ω is bounded. Let u be the distance function from $\mathbb{R}^{n-1} \setminus \Omega$. By [Maggi 2012, Remark 18.2], $\{u > \varepsilon\}$ is

an open bounded set of finite perimeter with $\partial^*\{u > \varepsilon\} =_{\mathcal{H}^{n-2}} \{u = \varepsilon\}$ for a.e. $\varepsilon > 0$. Moreover, if we set $f(x) = u(\boldsymbol{p}x)$, $x \in \mathbb{R}^n$, then $f : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function with $|\nabla f| = 1$ a.e. on $\Omega \times \mathbb{R}$, and $\{f = \varepsilon\} = \{u = \varepsilon\} \times \mathbb{R}$ for every $\varepsilon > 0$, so that, by the coarea formula for Lipschitz functions [Maggi 2012, Theorem 18.1],

$$\int_0^\infty \mathcal{H}^{n-1}\big(E^{(1)} \cap (\{u = \varepsilon\} \times \mathbb{R})\big) \, d\varepsilon = \int_{E^{(1)} \cap (\Omega \times \mathbb{R})} |\nabla f| \, d\mathcal{H}^n = \|v\|_{L^1(\Omega)} < \infty.$$

We may thus claim that, for a.e. $\varepsilon > 0$,

$$\mathcal{H}^{n-1}\big(E^{(1)} \cap (\partial^*\{u > \varepsilon\} \times \mathbb{R})\big) < \infty. \tag{A-3}$$

We now fix a sequence $\{\varepsilon_k\}_{k \in \mathbb{N}}$ such that $\varepsilon_k \to 0^+$ as $k \to \infty$, $\{u > \varepsilon_k\}$ is an open set of finite perimeter and $\varepsilon = \varepsilon_k$ satisfies (A-3) for every $k \in \mathbb{N}$. Now let $\{A_{k,i}\}_{i \in I_k}$ be the family of connected components of $\{u > \varepsilon_k\}$. Since $\partial A_{k,i} \subset \{u = \varepsilon_k\}$, and $\{u = \varepsilon_k\} =_{\mathcal{H}^{n-2}} \partial^*\{u > \varepsilon_k\}$ is $\mathcal{H}^{n-2}$-finite, we conclude by Federer's criterion that $A_{k,i}$ is of finite perimeter for every $k \in \mathbb{N}$ and $i \in I_k$. Let us now fix $z \in \Omega$, and let $k_0 \in \mathbb{N}$ be such that $z \in \{u > \varepsilon_k\}$ for every $k \geq k_0$. In this way, for every $k \geq k_0$, there exists $i_k(z) \in I_k$ such that $z \in A_{k,i_k(z)}$. We shall set

$$\Omega_k = A_{k,i_k(z)}.$$

By construction, each Ω_k is a bounded open connected set of finite perimeter, and $\Omega_k \subset \Omega_{k+1}$ for every $k \geq k_0$. Let us now prove $\Omega = \bigcup_{k \in \mathbb{N}} \Omega_k$. Indeed, let $y \in \Omega$, let $\gamma \in C^0([0, 1]; \Omega)$ with $\gamma(0) = z$ and $\gamma(1) = y$, and consider $K = \gamma([0, 1])$. Since K is compact, there exists $k_1 \in \mathbb{N}$ such that $K \subset \{u > \varepsilon_k\}$ for every $k \geq k_1$. Since K is connected and $\{z\} \subset K \cap \Omega_k$ for every $k \geq k_1$, we find that $K \subset \Omega_k$, and thus $y \in \Omega_k$, for every $k \geq k_1$. We now prove that E_k is a set of finite perimeter. Indeed, since $E_k = E \cap (\Omega_k \times \mathbb{R})$, we have $\partial^e E_k \subset [\partial^e E \cap (\overline{\Omega}_k \times \mathbb{R})] \cup [E^{(1)} \cap (\partial^e \Omega_k \times \mathbb{R})]$. Since Ω_k is compactly contained in Ω, we find $\mathcal{H}^{n-1}(\partial^e E \cap (\overline{\Omega}_k \times \mathbb{R})) \leq P(E; \Omega \times \mathbb{R}) < \infty$; thus, by taking (A-3) into account, we find $\mathcal{H}^{n-1}(\partial^e E_k) < \infty$, and thus that E_k is a set of finite perimeter thanks to Federer's criterion. By Proposition 3.2, $v_k \in BV(\mathbb{R}^{n-1})$ with $\mathcal{H}^{n-1}(\{v_k > 0\}) < \infty$, and $F[v_k]$ is a set of finite perimeter too. Since E_k is a v_k-distributed set of finite perimeter and $\partial^e \Omega_k$ is a countably $\mathcal{H}^{n-2}$-rectifiable set contained in $\{v_k^\wedge = 0\}$, by Proposition 3.8,

$$P(E_k; \partial^e \Omega_k \times \mathbb{R}) = P(F[v_k]; \partial^e \Omega_k \times \mathbb{R}).$$

Moreover, since $E_k = E \cap (\Omega_k \times \mathbb{R})$ and $F[v_k] = F[v] \cap (\Omega_k \times \mathbb{R})$,

$$P(E_k; \Omega_k^{(1)} \times \mathbb{R}) = P(E; \Omega_k^{(1)} \times \mathbb{R}), \quad P(F[v_k]; \Omega_k^{(1)} \times \mathbb{R}) = P(F[v]; \Omega_k^{(1)} \times \mathbb{R}).$$

Since $\Omega_k^{(0)} \times \mathbb{R} \subset E_k^{(0)} \cap F[v_k]^{(0)}$, we have proved (A-1) and (A-2). Finally, if $E \in \mathcal{M}_\Omega(v)$, then by (1-1) we have $P(E; \Omega_k \times \mathbb{R}) = P(F[v]; \Omega_k \times \mathbb{R})$, and thus, by (A-1) and (A-2), that $P(E_k) = P(F[v_k])$. $\quad\square$

Proof of Theorem B. Let $v \in BV(\Omega; [0, \infty))$ with $\mathcal{H}^{n-1}(\{v > 0\}) < \infty$, $D^s v \llcorner \{v^\wedge > 0\} = 0$ and $v^\wedge > 0$ $\mathcal{H}^{n-2}$-a.e. on Ω (so that $D^s v \llcorner \Omega = 0$). Let $E \in \mathcal{M}_\Omega(v)$, and assume for contradiction that $\mathcal{H}^n\big(E \Delta (t e_n + F[v])\big) > 0$ for every $t \in \mathbb{R}$. Let Ω_k be defined as in Lemma A.1, and let $v_k = 1_{\Omega_k} v$, $E_k = E \cap (\Omega_k \times \mathbb{R})$, so that $E_k \in \mathcal{M}(v_k)$ for every $k \in \mathbb{N}$. However, $\mathcal{H}^n\big(E_k \Delta (t e_n + F[v_k])\big) > 0$ for

every $t \in \mathbb{R}$ and for every k large enough. Thus, rigidity fails for v_k if k is large enough. By Theorem 1.11,

$$\{v_k^\wedge = 0\} \cup S_{v_k} \cup M_k \text{ essentially disconnects } \{v_k > 0\}, \tag{A-4}$$

where M_k is a concentration set for $D^c v_k$. Since $v_k^\wedge = 1_{\Omega_k^{(1)}} v^\wedge$ in Ω, $v^\wedge > 0$ $\mathcal{H}^{n-2}$-a.e. on Ω, and Ω_k is compactly contained in Ω, we find that

$$\{v_k^\wedge = 0\} = (\mathbb{R}^{n-1} \setminus \Omega_k^{(1)}) \cup (\{v^\wedge = 0\} \cap \Omega_k^{(1)}) =_{\mathcal{H}^{n-2}} \mathbb{R}^{n-1} \setminus \Omega_k^{(1)}.$$

Since $D^s v \llcorner \Omega = 0$, using Lemma 2.3 and (again) that Ω_k is compactly contained in Ω we find that

$$S_{v_k} \cap \Omega_k^{(1)} = S_v \cap \Omega_k^{(1)} =_{\mathcal{H}^{n-2}} S_v \cap (\Omega_k^{(1)} \setminus \Omega) = \varnothing.$$

Moreover, by Lemma 2.3, $D^c v_k = D^c v \llcorner \Omega_k^{(1)} = D^c v \llcorner (\Omega_k^{(1)} \setminus \Omega) = 0$, so that we may take $M_k = \varnothing$. Finally, $\{v_k > 0\}$ is $\mathcal{H}^{n-1}$-equivalent to Ω_k, and thus, by Remark 1.5, (A-4) can be equivalently rephrased as

$$(\mathbb{R}^{n-1} \setminus \Omega_k^{(1)}) \cup (S_{v_k} \setminus \Omega_k^{(1)}) \text{ essentially disconnects } \Omega_k. \tag{A-5}$$

In turn, this is equivalent to saying that Ω_k is not essentially connected. Since Ω_k is of finite perimeter, Ω_k is not indecomposable, by Remark 1.6. By [Ambrosio et al. 2001, Proposition 2], Ω_k is not connected. We have thus reached a contradiction. $\qquad \square$

Appendix B: A perimeter formula for vertically convex sets

We summarize here a perimeter formula for sets with segments as vertical sections that can be obtained as a consequence of Corollary 3.3 and Proposition 3.8, and that may be of independent interest.

Theorem B.1. *If $E = \{x \in \mathbb{R}^n : u_1(px) < qx < u_2(px)\}$ is a set of finite perimeter and volume defined by $u_1, u_2 : \mathbb{R}^{n-1} \to \mathbb{R}$ with $u_1 \leq u_2$ on $\mathbb{R}^{n-1}$, then u_1 and u_2 are approximately differentiable $\mathcal{H}^{n-1}$-a.e. on $\{u_2 > u_1\}$, and*

$$P(E) = \int_{\{v>0\}} \sqrt{1 + |\nabla u_1|^2} + \sqrt{1 + |\nabla u_2|^2} \, d\mathcal{H}^{n-1} + \int_{S_v \cup S_b} \min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\} \, d\mathcal{H}^{n-2}$$
$$+ |D^c u_1|^+(\{v^\wedge > 0\}) + |D^c u_2|^+(\{v^\wedge > 0\}),$$

where $v = u_2 - u_1$, $b = \frac{1}{2}(u_1 + u_2)$ and, for every Borel set $G \subset \mathbb{R}^{n-1}$, we set

$$|D^c u_i|^+(G) = \lim_{h \to \infty} |D^c(1_{\Sigma_h} u_i)|(G), \quad i = 1, 2, \tag{B-1}$$

where $\Sigma_h = \{\delta_h < v < L_h\}$ for sequences $\delta_h \to 0$ and $L_h \to \infty$ as $h \to \infty$ such that $\{v > \delta_h\}$ and $\{v < L_h\}$ are sets of finite perimeter. (Notice that $1_{\Sigma_h} u_i \in GBV(\mathbb{R}^{n-1})$ for $i = 1, 2$, so that $|D^c(1_{\Sigma_h} u_i)|$ are well-defined as Borel measures, and the right-hand side of (B-1) makes sense by monotonicity.)

Proof. By construction and by Theorem 1.7, if we set $v_h = 1_{\Sigma_h} v$ and $b_h = 1_{\Sigma_h} b$, then $v_h \in (BV \cap L^\infty)(\mathbb{R}^{n-1})$ and $b_h \in GBV(\mathbb{R}^{n-1})$ for every $h \in \mathbb{N}$, so that

$$1_{\Sigma_h} u_1 = b_h - \tfrac{1}{2} v_h \in GBV(\mathbb{R}^{n-1}), \quad 1_{\Sigma_h} u_2 = b_h + \tfrac{1}{2} v_h \in GBV(\mathbb{R}^{n-1}),$$

and, by Corollary 3.3, we find

$$P(E_h; G \times \mathbb{R}) = \int_{G \cap \{v_h > 0\}} \sqrt{1 + |\nabla b_h + \tfrac{1}{2}\nabla v_h|^2} + \sqrt{1 + |\nabla b_h - \tfrac{1}{2}\nabla v_h|^2} \, d\mathcal{H}^{n-1}$$
$$+ |D^c(b_h + \tfrac{1}{2}v_h)|(G \cap \{v_h^\wedge > 0\}) + |D^c(b_h - \tfrac{1}{2}v_h)|(G \cap \{v_h^\wedge > 0\})$$
$$+ \int_{G \cap (S_{v_h} \cup S_{b_h})} \min\{v_h^\vee + v_h^\wedge, \max\{[v_h], 2[b_h]\}\} \, d\mathcal{H}^{n-2}$$

for every Borel set $G \subset \mathbb{R}^{n-1}$, provided we set $E_h = W[v_h, b_h]$. Since $P(E; \Sigma_h^{(1)} \times \mathbb{R}) = P(E_h; \Sigma_h^{(1)} \times \mathbb{R})$, the above formula gives

$$P(E; \Sigma_h^{(1)} \times \mathbb{R}) = \int_{\Sigma_h} \sqrt{1 + |\nabla b + \tfrac{1}{2}\nabla v|^2} + \sqrt{1 + |\nabla b - \tfrac{1}{2}\nabla v|^2} \, d\mathcal{H}^{n-1}$$
$$+ \int_{\Sigma_h^{(1)} \cap (S_v \cup S_b)} \min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\} \, d\mathcal{H}^{n-2}$$
$$+ |D^c(b_h + \tfrac{1}{2}v_h)|(\{v^\wedge > 0\}) + |D^c(b_h - \tfrac{1}{2}v_h)|(\{v^\wedge > 0\}),$$

where we have also used that, for every $h \in \mathbb{N}$,

$$|D^c(b_h \pm \tfrac{1}{2}v_h)|(\Sigma_h^{(1)}) = |D^c(b_h \pm \tfrac{1}{2}v_h)|(\mathbb{R}^{n-1}) = |D^c(b_h \pm \tfrac{1}{2}v_h)|(\{v^\wedge > 0\}).$$

By monotonicity, and since $\bigcup_{h \in \mathbb{N}} \Sigma_h^{(1)} = \{v^\wedge > 0\} \cap \{v^\vee = \infty\} =_{\mathcal{H}^{n-2}} \{v^\wedge > 0\}$ — thanks to [Federer 1969, 4.5.9(3)] and since, by Proposition 3.2, $v \in BV(\mathbb{R}^{n-1})$ — we find that

$$P(E; \{v^\wedge > 0\} \times \mathbb{R})$$
$$= \int_{\{v > 0\}} \sqrt{1 + |\nabla u_1|^2} + \sqrt{1 + |\nabla u_2|^2} \, d\mathcal{H}^{n-1} + \int_{\{v^\wedge > 0\} \cap (S_v \cup S_b)} \min\{v^\vee + v^\wedge, \max\{[v], 2[b]\}\} \, d\mathcal{H}^{n-2}$$
$$+ |D^c u_1|^+(\{v^\wedge > 0\}) + |D^c u_2|^+(\{v^\wedge > 0\}).$$

At the same time, by Proposition 3.8, we have $P(E; \{v^\wedge = 0\} \times \mathbb{R}) = \int_{S_v \cap \{v^\wedge = 0\}} v^\vee \, d\mathcal{H}^{n-2}$. Adding up the last two identities we complete the proof of the formula for $P(E)$. $\qquad\square$

Acknowledgements

References

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Clarendon Press, New York, 2000. MR 2003a:49002 Zbl 0957.49001

[Ambrosio et al. 2001] L. Ambrosio, V. Caselles, S. Masnou, and J.-M. Morel, "Connected components of sets of finite perimeter and applications to image processing", *J. Eur. Math. Soc. (JEMS)* **3**:1 (2001), 39–92. MR 2002g:49072 Zbl 0981.49024

[Barchiesi et al. 2013] M. Barchiesi, F. Cagnetti, and N. Fusco, "Stability of the Steiner symmetrization of convex sets", *J. Eur. Math. Soc. (JEMS)* **15**:4 (2013), 1245–1278. MR 3055761 Zbl 1277.52012

[Cagnetti et al. 2013] F. Cagnetti, M. Colombo, G. D. Philippis, and F. Maggi, "Essential connectedness and the rigidity problem for Gaussian symmetrization", preprint, 2013. To appear in *J. European Math. Soc.* arXiv 1304.4527

[Chlebík et al. 2005] M. Chlebík, A. Cianchi, and N. Fusco, "The perimeter inequality under Steiner symmetrization: Cases of equality", *Ann. of Math.* (2) **162**:1 (2005), 525–555. MR 2006m:49032 Zbl 1087.28003

[De Giorgi 1958] E. De Giorgi, "Sulla proprietà isoperimetrica dell'ipersfera, nella classe degli insiemi aventi frontiera orientata di misura finita", *Atti Accad. Naz. Lincei. Mem. Cl. Sci. Fis. Mat. Nat. Sez. I* (8) **5** (1958), 33–44. MR 20 #4792 Zbl 0116.07901

[Dolzmann and Müller 1995] G. Dolzmann and S. Müller, "Microstructures with finite surface energy: The two-well problem", *Arch. Rational Mech. Anal.* **132**:2 (1995), 101–141. MR 97c:73018 Zbl 0846.73054

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Federer 1969] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften **153**, Springer, New York, 1969. MR 41 #1976 Zbl 0176.00801

[Giaquinta et al. 1998a] M. Giaquinta, G. Modica, and J. Souček, *Cartesian currents in the calculus of variations, I: Cartesian currents*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics **37**, Springer, Berlin, 1998. MR 2000b:49001a Zbl 0914.49001

[Giaquinta et al. 1998b] M. Giaquinta, G. Modica, and J. Souček, *Cartesian currents in the calculus of variations, II: Variational integrals*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics **38**, Springer, Berlin, 1998. MR 2000b:49001b Zbl 0914.49002

[Maggi 2012] F. Maggi, *Sets of finite perimeter and geometric variational problems: An introduction to geometric measure theory*, Cambridge Studies in Advanced Mathematics **135**, Cambridge University Press, 2012. MR 2976521 Zbl 1255.49074

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, 1995. MR 96h:28006 Zbl 0819.28004

FILIPPO CAGNETTI: f.cagnetti@sussex.ac.uk
Department of Mathematics, University of Sussex, Pevensey 2, Brighton, BN1 9QH, United Kingdom

MARIA COLOMBO: maria.colombo@sns.it
Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

GUIDO DE PHILIPPIS: guido.de.philippis@hcm.uni-bonn.de
Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, D-53115 Bonn, Germany

FRANCESCO MAGGI: maggi@math.utexas.edu
Department of Mathematics, The University of Texas at Austin, 2515 Speedway Stop C1200, Austin, TX 78712-1202, United States

msp

msp

SINGULAR BOHR–SOMMERFELD CONDITIONS
FOR 1D TOEPLITZ OPERATORS: HYPERBOLIC CASE

YOHANN LE FLOCH

We state the Bohr–Sommerfeld conditions around a singular value of hyperbolic type of the principal symbol of a selfadjoint semiclassical Toeplitz operator on a compact connected Riemann surface. These conditions allow the description of the spectrum of the operator in a fixed-size neighborhood of the singularity. We provide numerical computations for three examples, each associated with a different topology.

1. Introduction

Let M be a compact, connected Riemann surface with area form ω. Assume that M is endowed with a prequantum bundle L, that is, a Hermitian, holomorphic line bundle whose Chern connection has curvature $-i\omega$. Let K be another Hermitian holomorphic line bundle,[1] and define the quantum Hilbert space $\mathcal{H}_k$ as the space of holomorphic sections of $L^{\otimes k} \otimes K$, for every positive integer k. We consider (Berezin–)Toeplitz operators (see for instance [Boutet de Monvel and Guillemin 1981; Borthwick et al. 1998; Charles 2003a; Ma and Marinescu 2008] or the expository works [Ma 2010; Schlichenmaier 2010; Zelditch 2014]) acting on $\mathcal{H}_k$. The semiclassical limit corresponds to $k \to +\infty$.

The usual Bohr–Sommerfeld conditions, derived in [Charles 2006], describe the intersection of the spectrum of a selfadjoint Toeplitz operator and a neighborhood of any regular value of its principal symbol a_0 in terms of geometric quantities. More precisely, this intersection is the union of a finite number of families whose elements are, up to an error $O(k^{-2})$, the solutions of an equation of the form

$$c_0(\lambda) + k^{-1}(c_1(\lambda) + \epsilon\pi) \in 2\pi k^{-1}\mathbb{Z},$$

where

- $c_0(\lambda)$ is the holonomy associated with the parallel transport in L along a connected component of the level set $a_0^{-1}(\lambda)$,

- $c_1(\lambda)$ contains the integral of a differential form involving the subprincipal symbol of the operator,

- $\epsilon \in \{0, 1\}$ is an index associated with a half-forms structure.

[1]The reader must be warned that, in this work, the letter K does not refer to the canonical bundle unless explicitly stated otherwise.

Precise definitions of these quantities and a more explicit formulation of the Bohr–Sommerfeld rules can be found in Section 4F.

A natural question is whether one can write Bohr–Sommerfeld conditions near a singular value of the principal symbol. In the case of a nondegenerate singularity of elliptic type (a local extremum), it was answered positively in [Le Floch 2014], and the result is quite simple: roughly speaking, the singular Bohr–Sommerfeld conditions are nothing but the limit of the regular Bohr–Sommerfeld conditions when the energy tends to the singular value. The hyperbolic case (presence of saddle points) is much more difficult, because of the complicated topology of a neighborhood of the singular level. For instance, in the case of one hyperbolic point, the critical level looks like a figure eight, and crossing it has the effect of adding (or removing) one connected component from the regular level.

Let us mention that the case of Toeplitz operators is very close to the case of pseudodifferential operators. In this setting, the problem of describing the spectrum of a selfadjoint operator near a singular level of hyperbolic type was handled by Colin de Verdière and Parisse [1994a; 1994b; 1999]. In this article, we use analogous techniques to write hyperbolic Bohr–Sommerfeld conditions in the context of Toeplitz operators. The novelty is that they can be applied in this context.

1A. *Main result.* Let A_k be a selfadjoint Toeplitz operator on M; its normalized symbol $a_0 + \hbar a_1 + \cdots$ is real-valued. Assume that 0 is a critical value of the principal symbol a_0, that the level set $\Gamma_0 = a_0^{-1}(0)$ is connected and that every critical point contained in Γ_0 is nondegenerate and of hyperbolic type. Let $S = \{s_j\}_{1 \le j \le n}$ be the set of these critical points. Γ_0 is a compact graph embedded in M, and each of its vertices has degree 4 (this is a consequence of the usual Morse lemma, for instance). At each vertex s_j, we denote by e_m, $m = 1, 2, 3, 4$, the local edges, labeled with cyclic order $(1, 3, 2, 4)$ (with respect to the orientation of M near s_j) and such that e_1, e_2 (resp. e_3, e_4) correspond to the local unstable (resp. stable) manifolds. Cut $n + 1$ edges of Γ_0, each one corresponding to a cycle γ_i in a basis $(\gamma_1, \ldots, \gamma_{n+1})$ of $H_1(\Gamma_0, \mathbb{Z})$, in such a way that the remaining graph is a tree $\mathcal{T}$; usually $\mathcal{T}$ is called a spanning tree and the basis $(\gamma_1, \ldots, \gamma_{n+1})$ is called a fundamental cycle basis (see for instance [Berge 1973, pp. 25–26]). Our main result is the following:

Theorem (Theorem 6.1, Theorem 6.4). *Zero is an eigenvalue of A_k up to $O(k^{-\infty})$ if and only if the following system of $3n + 1$ linear equations with unknowns $(x_e \in \mathbb{C}_k)_{e \in \{\text{edges of } \mathcal{T}\}}$ (here $\mathbb{C}_k$ is the set of constant symbols; see Section 2A) has a nontrivial solution:*

(1) *If the edges (e_1, e_2, e_3, e_4) connect at s_j (with the same convention as before for their labeling), then*

$$\begin{pmatrix} x_{e_3} \\ x_{e_4} \end{pmatrix} = T_j \begin{pmatrix} x_{e_1} \\ x_{e_2} \end{pmatrix}.$$

(2) *If the edges α and β are the extremities of a cut cycle γ_i, then*

$$x_\alpha = \exp(ik\theta(\gamma_i, k))x_\beta,$$

where the following orientation is assumed: γ_i can be represented as a closed path starting on the edge α and ending on the edge β.

Moreover, T_j is a matrix depending only on a semiclassical invariant $\varepsilon_j(k)$ of the system at the singular point s_j (see (8)), and $\theta(\gamma, k)$ admits an asymptotic expansion in nonpositive powers of k. The first two terms of this expansion involve regularizations of the geometric invariants (actions and index) appearing in the usual Bohr–Sommerfeld conditions.

For spectral purposes, we use this theorem by replacing A_k by $A_k - E$ for E varying in a fixed-size neighborhood of the singular level. Away from the critical energy, we recover the regular Bohr–Sommerfeld conditions (see Section 6D).

This is very similar to the results of [Colin de Verdière and Parisse 1999], but the novelty lies in the framework that had to be set in order to extend their techniques to the Toeplitz setting (especially the sheaf-theoretic approach to the spectral theory of Toeplitz operators), and also in the geometric invariants that are specific to this context.

1B. *Structure of the article.* As said earlier, the case of Toeplitz operators is very close to the case of pseudodifferential operators; in mathematical terms, there is a microlocal equivalence between Toeplitz operators and pseudodifferential operators. When the phase space is the whole complex plane, this equivalence is realized by the Bargmann transform, and allows one to use some of the results obtained in the pseudodifferential setting. This is why the article is organized as follows: first, we discuss microlocal properties of the Bargmann transform. Then we introduce the sheaf of microlocal solutions of the equation $(A_k - E)u_k = 0$, explain its structure and recall the usual Bohr–Sommerfeld conditions. In Section 5, we construct a microlocal normal form for A_k near each critical point s_j, $1 \le j \le n$, on Bargmann spaces, and we use the properties of the Bargmann transform and the study of Colin de Verdière and Parisse [1994a] to describe the space of microlocal solutions of A_k near s_j. Finally, we adapt the reasoning of [Colin de Verdière and Parisse 1999; Colin de Verdière and Vũ Ngọc 2003] to obtain the singular Bohr–Sommerfeld conditions (in Section 6). We give numerical evidence in the last section.

2. Preliminaries and notation

2A. *Symbol classes.* We introduce rather standard symbol classes. Let d be a positive integer. For u in $\mathbb{C}^d \simeq \mathbb{R}^{2d}$, let $m(u) = (1 + \|u\|^2)^{1/2}$. For every integer j, we define the symbol class $\mathcal{S}_j^d$ as the set of sequences of functions in $\mathscr{C}^\infty(\mathbb{C}^d)$ which admit an asymptotic expansion of the form $a(\,\cdot\,, k) = \sum_{\ell \ge 0} k^{-\ell} a_\ell$ in the sense that

- For all $\ell \in \mathbb{N}$ and all $\alpha, \beta \in \mathbb{N}^{2d}$, there exists $C_{\ell, \alpha, \beta} > 0$ such that $|\partial_z^\alpha \partial_{\bar{z}}^\beta a_\ell| \le C_{\ell, \alpha, \beta} m^j$.

- For all $L \in \mathbb{N}^*$ and all $\alpha, \beta \in \mathbb{N}^{2d}$, there exists $C_{L, \alpha, \beta} > 0$ such that

$$\left| \partial_z^\alpha \partial_{\bar{z}}^\beta \left(a - \sum_{\ell=0}^{L-1} k^{-\ell} a_\ell \right) \right| \le C_{L, \alpha, \beta} k^{-L} m^j.$$

We set $\mathcal{S}^d = \bigcup_{j \in \mathbb{Z}} \mathcal{S}_j^d$. If, in the definition of $\mathcal{S}_0^1$, we only consider symbols independent of z, we obtain the class $\mathbb{C}_k$ of constant symbols.

2B. *Function spaces.* Using standard notation, we denote by $\mathscr{S}(\mathbb{R})$ the Schwartz space of functions $f \in \mathscr{C}^\infty(\mathbb{R})$ such that $\sup_{t \in \mathbb{R}} |t^j f^{(p)}(t)| < +\infty$ for all $j, p \in \mathbb{N}$, by $\mathscr{D}'(\mathbb{R})$ the space of distributions on $\mathbb{R}$, and by $\mathscr{S}'(\mathbb{R}) \subset \mathscr{D}'(\mathbb{R})$ the space of tempered distributions on $\mathbb{R}$ (the dual space of $\mathscr{S}(\mathbb{R})$). We recall that

$$\mathscr{S}(\mathbb{R}) = \bigcap_{j \in \mathbb{N}} \mathscr{S}_j(\mathbb{R}),$$

where $\mathscr{S}_j(\mathbb{R})$ is the space of functions f in $\mathscr{C}^j(\mathbb{R})$ with $\|f\|_{\mathscr{S}_j}$ finite, with

$$\|f\|_{\mathscr{S}_j} = \max_{0 \le p \le j} \left(\sup_{t \in \mathbb{R}} \left| (1 + t^2)^{(j-p)/2} f^{(p)}(t) \right| \right).$$

The topology of $\mathscr{S}(\mathbb{R})$ is defined by the countable family of seminorms $\| \cdot \|_{\mathscr{S}_j}$, $j \in \mathbb{N}$.

We recall the definition of Bargmann spaces [Bargmann 1961; 1967], which are spaces of square-integrable functions with respect to a Gaussian weight: for $k \in \mathbb{N}^*$,

$$\mathscr{B}_k = \left\{ f \psi^k \ \middle| \ f : \mathbb{C} \mapsto \mathbb{C} \text{ holomorphic}, \int_{\mathbb{R}^2} |f(z)|^2 \exp(-k|z|^2) \, d\lambda(z) < +\infty \right\}$$

with $\psi : \mathbb{C} \to \mathbb{C}, z \mapsto \exp\!\left(-\frac{1}{2}|z|^2\right)$, $\psi^k : \mathbb{C} \to \mathbb{C}^{\otimes k}$ its k-th tensor power, and λ the Lebesgue measure on $\mathbb{R}^2$. We denote by $\| \cdot \|_{\mathscr{B}_k}$ the naturally associated L^2-norm:

$$\|f \psi^k\|_{\mathscr{B}_k} = \left(\int_{\mathbb{R}^2} |f(z)|^2 \exp(-k|z|^2) \, d\lambda(z) \right)^{1/2}.$$

Of course, this norm is still defined for elements of the form $f \psi^k$ satisfying the integrability condition with f not necessarily holomorphic; when this is the case, we denote it by $\|f \psi^k\|_{L^2, \exp}$. Furthermore, we introduce the subspace

$$\mathfrak{S}_k = \left\{ \varphi \in \mathscr{B}_k \ \middle| \ \forall j \in \mathbb{N}, \ \sup_{z \in \mathbb{C}} \left(|\varphi(z)|(1 + |z|^2)^{j/2} \right) < +\infty \right\} \tag{1}$$

of $\mathscr{B}_k$, with topology induced by the obvious associated family of seminorms. It is the analogue of the Schwartz space on the Bargmann side; see Section 3A for a more precise statement.

2C. *Weyl quantization and pseudodifferential operators.* We briefly recall some standard notation and properties of the theory of pseudodifferential operators (for details, see e.g. [Colin de Verdière 2009; Dimassi and Sjöstrand 1999; Zworski 2012]), replacing the usual small parameter $\hbar$ by k^{-1}, because this is all we need in the rest of the paper.

2C1. *Pseudodifferential operators.* A pseudodifferential operator in one degree of freedom is an operator (possibly unbounded) acting on $L^2(\mathbb{R})$ which is the Weyl quantization of a symbol $a(\,\cdot\,, k) \in \mathscr{S}^1$, seen as a sequence of functions defined on the cotangent space $T^*\mathbb{R} \simeq \mathbb{R}^2$; more precisely,

$$(\mathrm{Op}_k^W(a)u)(x) = \frac{k}{2\pi} \int_{\mathbb{R}^2} \exp(ik(x - y)\xi) a\!\left(\frac{x+y}{2}, \xi, k \right) u(y) \, dy \, d\xi.$$

The leading term a_0 in the asymptotic expansion of $a(\,\cdot\,, k)$ is the principal symbol of $A_k = \mathrm{Op}_k^W(a)$. A_k is said to be *elliptic* at $(x_0, \xi_0) \in T^*\mathbb{R}$ if $a_0(x_0, \xi_0) \ne 0$.

2C2. *Wavefront set.*

Definition 2.1. A sequence u_k of elements of $\mathscr{D}'(\mathbb{R})$ is said to be *admissible* if, for any pseudodifferential operator P_k whose symbol is compactly supported, there exists an integer $N \in \mathbb{Z}$ such that $\|P_k u_k\|_{L^2(\mathbb{R})} = O(k^N)$.

We recall the standard definition of the wavefront set $\mathrm{WF}(u_k)$ of an admissible sequence of distributions.

Definition 2.2. Let u_k be an admissible sequence in $\mathscr{D}'(\mathbb{R})$. A point (x_0, ξ_0) does not belong to $\mathrm{WF}(u_k)$ if and only if there exists a pseudodifferential operator P_k, elliptic at (x_0, ξ_0), such that $\|P_k u_k\|_{L^2(\mathbb{R})} = O(k^{-\infty})$.

One can refine these definitions in the case where u_k belong to $\mathscr{S}(\mathbb{R})$.

Definition 2.3. A sequence $(u_k)_{k \geq 1}$ of elements of $\mathscr{S}(\mathbb{R})$ is said to be

- $\mathscr{S}$-*admissible* if there exists N in $\mathbb{Z}$ such that every Schwartz seminorm of u_k is $O(k^N)$,

- $\mathscr{S}$-*negligible* if every Schwartz seminorm of u_k is $O(k^{-\infty})$. We write $u_k = O_{\mathscr{S}}(k^{-\infty})$.

Now, instead of using the L^2-norm in Definition 2.2, one can actually consider the seminorms $\| \cdot \|_{\mathscr{S}_j}$.

Lemma 2.4. *A point (x_0, ξ_0) does not belong to $\mathrm{WF}(u_k)$ if and only if there exists a pseudodifferential operator P_k, elliptic at (x_0, ξ_0), such that $P_k u_k = O_{\mathscr{S}}(k^{-\infty})$.*

Proof. The sufficient condition comes from the previous definition, so we only prove the necessary condition. We only adapt a standard argument used when one wants to deal with $\mathscr{C}^j$-norms (see [Robert 1987, Proposition IV-8]). Assume that (x_0, ξ_0) does not belong to $\mathrm{WF}(u_k)$; there exists a pseudodifferential operator P_k, elliptic at (x_0, ξ_0), such that $\|P_k u_k\|_{L^2(\mathbb{R})} = O(k^{-\infty})$. Consider a compactly supported smooth function χ equal to one in a neighborhood of (x_0, ξ_0), and set $Q_k = \mathrm{Op}^{\mathrm{W}}(\chi) P_k$. For every $R \in \mathbb{R}[X]$ and every integer $j > 0$,

$$k^{-j} \frac{d^j}{dx^j} R \, \mathrm{Op}^{\mathrm{W}}(\chi)$$

is a pseudodifferential operator of order 0, hence bounded $L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by a constant $C > 0$ (by the Calderón–Vaillancourt theorem; see [Robert 1987, Theorem II-36] or [Dimassi and Sjöstrand 1999, Theorem 7.11]). Thus, one has

$$\left\| k^{-j} \frac{d^j}{dx^j} R Q_k u_k \right\|_{L^2(\mathbb{R})} \leq C \|P_k u_k\|_{L^2(\mathbb{R})} = O(k^{-\infty}).$$

Hence, $\|R Q_k u_k\|_{H^s(\mathbb{R})} = O(k^{-\infty})$ for every integer $s > 0$, where we recall that the Sobolev space $H^s(\mathbb{R})$ is the subspace of $L^2(\mathbb{R})$ whose elements have their s first derivatives in $L^2(\mathbb{R})$; Sobolev injections then yield that every $\mathscr{C}^j$-norm of $R Q_k u_k$ is $O(k^{-\infty})$. Since this holds for every polynomial R, we obtain the result. $\qquad\square$

2D. *Geometric quantization and Toeplitz operators.* We also recall the standard definitions and notation in the Toeplitz setting. Unless otherwise mentioned, "smooth" will always mean $\mathscr{C}^\infty$, and a section of a line bundle will always be assumed to be smooth. The space of sections of a bundle $E \to M$ will be denoted by $\mathscr{C}^\infty(M, E)$. Let M be a connected, compact Kähler manifold, with fundamental 2-form $\omega \in \Omega^2(M, \mathbb{R})$. Assume M is endowed with a prequantum bundle $L \to M$, that is, a Hermitian holomorphic line bundle whose Chern connection ∇ has curvature $-i\omega$. Let $K \to M$ be a Hermitian holomorphic line bundle. For every positive integer k, define the quantum space $\mathscr{H}_k$ as

$$\mathscr{H}_k = H^0(M, L^k \otimes K) = \{\text{holomorphic sections of } L^k \otimes K\}.$$

The space $\mathscr{H}_k$ is a subspace of the space $L^2(M, L^k \otimes K)$ of sections of finite L^2-norm, where the scalar product is given by

$$\langle \varphi, \psi \rangle = \int_M h_k(\varphi, \psi) \mu_M,$$

with h_k the Hermitian product on $L^k \otimes K$ induced by those of L and K, and μ_M the Liouville measure on M. Since M is compact, $\mathscr{H}_k$ is finite-dimensional, and is thus given a Hilbert space structure with this scalar product.

2D1. *Admissible and negligible sequences.* Let $(s_k)_{k \geq 1}$ be a sequence such that, for each k, s_k belongs to $\mathscr{C}^\infty(M, L^k \otimes K)$. We say that $(s_k)_{k \geq 1}$ is

- *admissible* if for every positive integer ℓ, for all vector fields $X_1, \ldots, X_\ell$ on M, and for every compact set $C \subset M$, there exist a constant $c > 0$ and an integer N (depending on $X_1, \ldots, X_\ell$ and C) such that

$$\|\nabla_{X_1} \cdots \nabla_{X_\ell} s_k(p)\| \leq ck^N \quad \text{for all } p \in C;$$

- *negligible* if for all positive integers ℓ and N, for all vector fields $X_1, \ldots, X_\ell$ on M, and for every compact set $C \subset M$, there exists a constant $c > 0$ (depending on $X_1, \ldots, X_\ell$, C and N) such that

$$\|\nabla_{X_1} \cdots \nabla_{X_\ell} s_k(p)\| \leq ck^{-N} \quad \text{for all } p \in C.$$

We say that $(s_k)_{k \geq 1}$ is *negligible* over an open set $U \subset M$ if the previous estimates hold for every compact subset of U. The *microsupport* $\text{MS}(s_k)$ of an admissible sequence $(s_k)_{k \geq 1}$ is the complement of the set of points of M which admit a neighborhood where $(s_k)_{k \geq 1}$ is negligible. Finally, we say that two admissible sequences $(t_k)_{k \geq 1}$ and $(s_k)_{k \geq 1}$ are *microlocally equal* on an open set U if $\text{MS}(t_k - s_k) \cap U = \varnothing$; unless explicitly stated otherwise, the symbol $\sim$ will indicate microlocal equivalence.

2D2. *Toeplitz operators.* Let Π_k be the orthogonal projector of $L^2(M, L^k \otimes K)$ onto $\mathscr{H}_k$. A *Toeplitz operator* is any sequence $(T_k : \mathscr{H}_k \to \mathscr{H}_k)_{k \geq 1}$ of operators of the form

$$T_k = \Pi_k M_{f(\,\cdot\,,k)} + R_k,$$

where $f(\,\cdot\,, k)$ is a sequence in $\mathscr{C}^\infty(M)$ with an asymptotic expansion $f(\,\cdot\,, k) = \sum_{\ell \geq 0} k^{-\ell} f_\ell$ for the $\mathscr{C}^\infty$ topology, $M_{f(\,\cdot\,,k)}$ is the operator of multiplication by $f(\,\cdot\,, k)$, and R_k is an operator acting on $\mathscr{H}_k$

with $\|R_k\| = O(k^{-\infty})$. Let $\mathcal{T}$ be the set of Toeplitz operators, and define the contravariant symbol map

$$\sigma_{\mathrm{cont}} : \mathcal{T} \to \mathscr{C}^\infty(M)[\![\hbar]\!]$$

sending T_k into the formal series $\sum_{\ell \geq 0} \hbar^\ell f_\ell$. We will mainly work with the *normalized symbol*

$$\sigma_{\mathrm{norm}} = \Big(\mathrm{Id} + \frac{\hbar}{2}\Delta\Big)\sigma_{\mathrm{cont}},$$

where $\Delta = \partial^*\partial$ is the holomorphic Laplacian acting on $\mathscr{C}^\infty(M)$; unless otherwise mentioned, when we talk about a subprincipal symbol, this refers to the normalized symbol.

We can define the notions of admissibility, negligibility, microsupport and microlocal equivalence for Toeplitz operators, using the fact that their Schwartz kernels are sequences of sections of some line bundle (see [Charles 2006, Equation (4.1)] for a more precise statement).

2D3. *The case of the complex plane.* Let us briefly recall how to adapt the previous constructions to the case of the whole complex plane. We consider the Kähler manifold $\mathbb{C} \simeq \mathbb{R}^2$ with coordinates (x, ξ), standard complex structure and symplectic form $\omega_0 = d\xi \wedge dx$. Let $L_0 = \mathbb{R}^2 \times \mathbb{C} \to \mathbb{R}^2$ be the trivial fiber bundle with standard Hermitian metric h_0 and connection ∇^0 with 1-form $(1/i)\alpha$, where $\alpha_u(v) = \frac{1}{2}\omega_0(u, v)$; endow L_0 with the unique holomorphic structure compatible with h_0 and ∇^0. For every positive integer k, the quantum space at order k is

$$\mathcal{H}_k^0 = H^0(\mathbb{R}^2, L_0^k) \cap L^2(\mathbb{R}^2, L_0^k),$$

and it turns out that $\mathcal{H}_k^0 = \mathcal{B}_k$ (see Section 2B for the definition of $\mathcal{B}_k$); indeed, if we choose the holomorphic coordinate $z = \frac{1}{\sqrt{2}}(x - i\xi)$, then a section φ of $L_0^k \to \mathbb{R}^2$ is holomorphic if and only if

$$\partial_{\bar{z}}\varphi + \frac{kz}{2}\varphi = 0.$$

Hence, for $\psi : \mathbb{C} \to \mathbb{C}$, $z \mapsto \exp\big(-\frac{1}{2}|z|^2\big)$, the section ψ^k is a nonvanishing element of $H^0(\mathbb{R}^2, L_0^k)$, and any other holomorphic section is of the form $f\psi^k$, where f is a holomorphic function.

One can define the algebra of Toeplitz operators and the various symbols in a similar way as in the compact case; see [Le Floch 2014] for details. We will call $\mathcal{T}_j$ the class of Toeplitz operators with symbol in $\mathscr{S}_j^1$. In what follows, Π_k^0 will denote the orthogonal projector of $L^2(\mathbb{R}^2, L_0^k)$ onto $\mathcal{H}_k^0$, and we define the Toeplitz operator $\mathrm{Op}(f \cdot, k) = \Pi_k^0 M_{f(\cdot, k)}$ for $f(\cdot, k)$ in $\mathscr{S}_j^1$.

Let us give more details about the microsupport in this setting. We start by recalling the following inequality in Bargmann spaces [Bargmann 1961, Equation (1.7)].

Lemma 2.5. *Let $\phi_k \in \mathcal{B}_k$. Then, for every complex variable z,*

$$|\phi_k(z)| \leq \Big(\frac{k}{2\pi}\Big)^{1/2}\|\phi_k\|_{\mathcal{B}_k}.$$

Similarly, for all vector fields $X_1, \ldots, X_p$ on $\mathbb{C}$, there exists a polynomial $P \in \mathbb{R}[x_1, x_2]$ with positive values such that, for every $z \in \mathbb{C}$,

$$|(\nabla_{X_1} \cdots \nabla_{X_p}\phi_k)(z)| \leq P(|z|, k)^{1/2}\|\phi_k\|_{\mathcal{B}_k}.$$

Proof. The first claim is proved in [Bargmann 1961] in the case $k = 1$; the general case then comes from a change of variables. The second claim can be proved in the same way. $\qquad\square$

Lemma 2.6. *Let u_k be a sequence of elements of $\mathcal{B}_k$ and Ω a bounded open subset of $\mathbb{C}$. Assume that $\|u_k\|_{L^2(\Omega),\exp} = O(k^{-\infty})$; then, for any compact subset K of Ω, u_k and all its covariant derivatives are uniformly $O(k^{-\infty})$ on K.*

Proof. Choose a compactly supported smooth function η which is positive, vanishing outside Ω and with constant value 1 on K, and set $v_k = \mathrm{Op}(\eta)u_k$. One has

$$\|v_k\|_{\mathcal{B}_k} = \|\Pi_k^0 \eta u_k\|_{\mathcal{B}_k} \leq \|\eta u_k\|_{L^2,\exp} \leq \|u_k\|_{L^2(\Omega),\exp}$$

since Π_k^0 is continuous $L^2(\mathbb{R}^2, L_0^k) \to L^2(\mathbb{R}^2, L_0^k)$ with norm smaller than 1. Hence, $\|v_k\|_{\mathcal{B}_k} = O(k^{-\infty})$. By Lemma 2.5, this implies that v_k and its covariant derivatives are uniformly $O(k^{-\infty})$ on K; since $u_k = v_k + O(k^{-\infty})$ on K, the same holds for u_k. $\qquad\square$

Lemma 2.7. *Let $(u_k)_{k \geq 1}$ be an admissible sequence of elements of $\mathcal{B}_k$ and $z_0 \in \mathbb{C}$. Then $z_0 \notin \mathrm{MS}(u_k)$ if and only if there exists a Toeplitz operator $T_k \in \mathcal{T}_0$, elliptic at z_0, such that $\|T_k u_k\|_{\mathcal{B}_k} = O(k^{-\infty})$.*

Proof. Assume that $z_0 \notin \mathrm{MS}(u_k)$. There exists a neighborhood $\mathcal{U}$ of z_0 such that u_k is negligible on $\mathcal{U}$. Choose a compactly supported function $\chi \in \mathscr{C}^\infty(\mathbb{C}, \mathbb{R})$ with support K contained in $\mathcal{U}$ and such that $\chi(z_0) = 1$, and set $T_k = \mathrm{Op}(\chi)$. One has, for $z_1 \in \mathbb{C}$,

$$(T_k u_k)(z_1) = \frac{k}{2\pi} \int_K \exp\left(-\frac{k}{2}(|z_1|^2 + |z_2|^2 - 2z_1 \bar{z}_2)\right) \chi(z_2) u_k(z_2)\, d\mu(z_2),$$

which gives

$$|(T_k u_k)(z_1)| \leq \frac{k}{2\pi} \sup_K |u_k| \int_K \exp\left(-\frac{k}{2}|z_1 - z_2|^2\right) d\mu(z_2).$$

This allows to estimate the norm of $T_k u_k$:

$$\|T_k u_k\|_{\mathcal{B}_k}^2 \leq \left(\frac{k}{2\pi}\right)^2 \left(\sup_K |u_k|\right)^2 \int_\mathbb{C} \int_K \exp(-k|z_1 - z_2|^2)\, d\mu(z_1)\, d\mu(z_2).$$

Hence

$$\|T_k u_k\|_{\mathcal{B}_k}^2 \leq \left(\frac{k}{2\pi}\right)^2 \left(\sup_K |u_k|\right)^2 \mu(K) \int_\mathbb{C} \exp(-k|z_1|^2)\, d\mu(z_1),$$

and the necessary condition is proved since the integral is $O(k^{-1/2})$.

Conversely, assume that there exists a Toeplitz operator $T_k \in \mathcal{T}_0$, elliptic at z_0, such that $\|T_k u_k\|_{\mathcal{B}_k} = O(k^{-\infty})$. There exists a neighborhood of z_0 where T_k is elliptic. Hence, by symbolic calculus, we can find a Toeplitz operator $S_k \in \mathcal{T}_0$ such that $S_k T_k \sim \Pi_k^0$ near (z_0, z_0). Thus, there exists a neighborhood Ω of z_0 such that $S_k T_k u_k \sim u_k$ on Ω; this implies that $\|S_k T_k u_k\|_{L^2(\Omega)} = \|u_k\|_{L^2(\Omega)} + O(k^{-\infty})$. But, since S_k is bounded $\mathcal{B}_k \to \mathcal{B}_k$ by a constant $C > 0$ which does not depend on k, one has $\|S_k T_k u_k\|_{L^2(\Omega)} \leq C \|T_k u_k\|_{\mathcal{B}_k}$; this yields that $\|u_k\|_{L^2(\Omega)}$ is $O(k^{-\infty})$. Lemma 2.6 then gives the negligibility of u_k on Ω. $\qquad\square$

Definition 2.8. A sequence $(u_k)_{k \geq 1}$ of elements of $\mathfrak{S}_k$ is said to be

- $\mathfrak{S}_k$-*admissible* if there exists N in $\mathbb{Z}$ such that every $\mathfrak{S}_k$ seminorm of u_k is $O(k^N)$;

- $\mathfrak{S}_k$-*negligible* if every $\mathfrak{S}_k$ seminorm of u_k is $O(k^{-\infty})$. We write $u_k = O_{\mathfrak{S}_k}(k^{-\infty})$.

Lemma 2.9. *Let $(u_k)_{k\geq 1}$ be an admissible sequence of elements of $\mathcal{B}_k$ and $z_0 \in \mathbb{C}$. Then $z_0 \notin \mathrm{MS}(u_k)$ if and only if there exists a Toeplitz operator $T_k \in \mathcal{T}_0$, elliptic at z_0, such that $T_k u_k = O_{\mathfrak{S}_k}(k^{-\infty})$.*

Proof. The proof is nearly the same as the one of Lemma 2.4. One can show that if $z_0 \notin \mathrm{MS}(u_k)$, there exists a Toeplitz operator $T_k \in \mathcal{T}_0$, elliptic at z_0, such that for every polynomial function $P(z)$ of z only, $\sup_{z\in\mathbb{C}} |P(z)(T_k u_k)(z)| = O(k^{-\infty})$, using the fact that the multiplication by $P(z)$ is a Toeplitz operator. $\square$

3. The Bargmann transform

3A. *Definition and first properties.* The Bargmann transform is the unitary operator $B_k : L^2(\mathbb{R}) \to \mathcal{B}_k$ defined by

$$(B_k f)(z) = \left(\left(\frac{k}{\pi}\right)^{1/4} \int_{\mathbb{R}} \exp\left(k\left(-\tfrac{1}{2}(z^2 + t^2) + \sqrt{2}zt\right)\right) f(t)\, dt \right) \psi^k(z).$$

We claimed earlier that the subspace $\mathfrak{S}_k$ of $\mathcal{B}_k$ defined in (1) is the analogue of the Schwartz space on the Bargmann side. The case $k = 1$ is treated by the following theorem.

Theorem 3.1 [Bargmann 1967, Theorem 1.7]. *The Bargmann transform B_1 is a bijective, bicontinuous mapping between $\mathscr{S}(\mathbb{R})$ and $\mathfrak{S}_1$.*

This allows us to handle the general case.

Proposition 3.2. *The Bargmann transform B_k is a bijection between $\mathscr{S}(\mathbb{R})$ and $\mathfrak{S}_k$.*

Proof. If f belongs to $\mathscr{S}(\mathbb{R})$, one has, for z in $\mathbb{C}$,

$$(B_k f)(z) = \left(\frac{k}{\pi}\right)^{1/4} \int_{\mathbb{R}} \exp\left(k\left(-\tfrac{1}{2}(z^2 + t^2) + \sqrt{2}zt\right)\right) f(t)\, dt;$$

introducing the variables u and w such that $z = k^{-1/2}w$ and $t = k^{-1/2}u$, this reads

$$(B_k f)(z) = (k\pi)^{-1/4} \int_{\mathbb{R}} \exp\left(-\tfrac{1}{2}(w^2 + u^2) + \sqrt{2}wu\right) f(k^{-1/2}u)\, du.$$

Hence, we have $(B_k f)(z) = (k\pi)^{-1/4}(B_1 g)(k^{1/2}z)$, where $g(t) = f(k^{-1/2}t)$. Obviously, the function g belongs to $\mathscr{S}(\mathbb{R})$; thus, by the previous theorem, $B_1 g$ belongs to $\mathfrak{S}_1$. Hence, for $j \in \mathbb{N}$, there exists a constant $C_j > 0$ such that, for every complex variable w,

$$\left| (B_1 g)(w) \exp\left(-\tfrac{1}{2}|w|^2\right) \right| \leq C_j (1 + |w|^2)^{-j/2}.$$

This implies that, for every z in $\mathbb{C}$,

$$\left| (B_k f)(z) \exp\left(-\frac{k}{2}|z|^2\right) \right| \leq C_j k^{-j/2}(1 + k|z|^2)^{-j/2},$$

and since $k \geq 1$, this yields

$$\left| (B_k f)(z) \exp\left(-\frac{k}{2}|z|^2\right) \right| \leq C_j (1 + |z|^2)^{-j},$$

which means that $B_k f$ belongs to $\mathfrak{S}_k$. The converse is proved in the same way, using the explicit form of the inverse mapping:

$$(B_k^* g)(t) = \left(\frac{k}{\pi}\right)^{1/4} \int_{\mathbb{R}} \exp\left(k\left(-\tfrac{1}{2}(\bar{z}^2 + t^2) + \sqrt{2}\bar{z}t - |z|^2\right)\right) g(z)\, d\mu(z)$$

for g in $\mathfrak{S}_k$ and $t \in \mathbb{R}$. $\square$

3B. *Action on Toeplitz operators.* The Bargmann transform has the good property to conjugate a Toeplitz operator with symbol defined on $\mathbb{C}$ (thus acting on the spaces $\mathscr{B}_k$) to a pseudodifferential operator with symbol defined on $T^*\mathbb{R}$ (thus acting on $L^2(\mathbb{R})$), and conversely.

Lemma 3.3. *Let T_k be a Toeplitz operator in the class $\mathscr{T}_j$, with contravariant symbol $\sigma_{\mathrm{cont}}(T_k) = f(\,\cdot\,, \hbar)$. Then $B_k^* T_k B_k$ is a pseudodifferential operator with Weyl symbol*

$$\sigma^W(B_k^* T_k B_k)(x, \xi) = I(f(\,\cdot\,, \hbar))(x, \xi) = \frac{1}{\pi \hbar} \int_{\mathbb{C}} \exp(-2\hbar^{-1}|w|^2) f(w + z, \hbar)\, d\lambda(w),$$

where $z = \frac{1}{\sqrt{2}}(x - i\xi)$. The map I is continuous $\mathscr{S}_j \to \mathscr{S}_j$. Moreover, for any $f(\,\cdot\,, \hbar) \in \mathscr{S}_j$ and all $p \geq 1$,

$$I(f(\,\cdot\,, \hbar)) = \sum_{j=0}^{p-1} \left(\frac{\hbar}{2}\right)^j \frac{\Delta^j f(\,\cdot\,, \hbar)}{j!} + h^p R_p(f(\,\cdot\,, \hbar)), \tag{2}$$

where R_p is a continuous map from $\mathscr{S}_j$ to $\mathscr{S}_j$.

Proof. Thanks to [Charles and Vũ Ngọc 2008, Theorem 5.2], we know that the result holds when $T_k = \Pi_k^0 f \Pi_k^0$, f being a bounded function on $\mathbb{C}$ not depending on k. Now, using the stationary phase method, one can prove that the map I is continuous $\mathscr{S}_j \to \mathscr{S}_j$ with the asymptotic expansion (2), and conclude by a density argument. $\square$

3C. *Microlocalization and Bargmann transform.*

Lemma 3.4. (1) *B_k maps $\mathscr{S}$-admissible functions to $\mathfrak{S}_k$-admissible sections, and B_k^* maps $\mathfrak{S}_k$-admissible sections to $\mathscr{S}$-admissible functions.*

(2) *B_k maps $O_{\mathscr{S}}(k^{-\infty})$ into $O_{\mathfrak{S}_k}(k^{-\infty})$, and B_k^* maps $O_{\mathfrak{S}_k}(k^{-\infty})$ into $O_{\mathscr{S}}(k^{-\infty})$.*

Proof. These results are proved by performing a change of variables, as in Proposition 3.2. $\square$

We can now prove the link between the wavefront set and the microsupport via the Bargmann transform.

Proposition 3.5. *Let u_k be an admissible sequence of elements of $\mathscr{S}(\mathbb{R})$. Then $(x_0, \xi_0) \notin \mathrm{WF}(u_k)$ if and only if $z_0 = \frac{1}{\sqrt{2}}(x_0 - i\xi_0) \notin \mathrm{MS}(B_k u_k)$.*

Proof. Assume that $z_0 = \frac{1}{\sqrt{2}}(x_0 - i\xi_0)$ does not belong to $\mathrm{MS}(B_k u_k)$; by Lemma 2.9, there exists a Toeplitz operator T_k, elliptic at z_0, such that $T_k B_k u_k \psi^k = O_{\mathfrak{S}_k}(k^{-\infty})$. Thanks to Lemma 3.3, $P_k = B_k^* T_k B_k$ is a pseudodifferential operator elliptic at (x_0, ξ_0). Furthermore, thanks to Lemma 3.4, $P_k u_k = B_k^* T_k B_k u_k \psi_k = O_{\mathscr{S}}(k^{-\infty})$; we conclude by Lemma 2.4. The proof of the converse follows the same steps. $\square$

4. The sheaf of microlocal solutions

In this section, T_k is a selfadjoint Toeplitz operator on M with normalized symbol $f(\,\cdot\,, \hbar) = \sum_{\ell \geq 0} \hbar^\ell f^\ell$. Following [Vũ Ngọc 1998; 2000], we introduce the sheaf of microlocal solutions of the equation $T_k \psi_k = 0$.

Let us recall the motivation for considering microlocal solutions: roughly speaking, they allow to split the eigenvalue equation $T_k \psi_k = \lambda \psi_k$ into several local problems, the Bohr–Sommerfeld rules being a necessary and sufficient condition to glue together the solutions to these problems in order to obtain a global approximate solution to this equation. For the sake of brevity, we begin with the case $\lambda = 0$, and we introduce a spectral parameter only in Section 4F.

4A. *Microlocal solutions.* For an open subset U of M, we call a sequence of sections $\psi_k \in \mathscr{C}^\infty(U, L^k \otimes K)$ a local state over U.

Definition 4.1. We say that a local state ψ_k is a microlocal solution of

$$T_k \psi_k = 0 \tag{3}$$

on U if it is admissible and, for every $x \in U$, there exists a function $\chi \in \mathscr{C}^\infty(M)$ with support contained in U, equal to 1 in a neighborhood of x and such that

$$\Pi_k(\chi \psi_k) = \psi_k + O(k^{-\infty}), \quad T_k(\Pi_k(\chi \psi_k)) = O(k^{-\infty})$$

on a neighborhood of x.

One can show that if $\psi_k \in \mathscr{H}_k$ is admissible and satisfies $T_k \psi_k = 0$, then the restriction of ψ_k to U is a microlocal solution of (3) on U. Moreover, the set $S(U)$ of microlocal solutions of this equation on U is a $\mathbb{C}_k$-module containing the set of negligible local states as a submodule (let us recall that $\mathbb{C}_k$ is the set of constant symbols; see Section 2A). We denote by $\mathrm{Sol}(U)$ the module obtained by taking the quotient of $S(U)$ by the negligible local states; the notation $[\psi_k]$ will stand for the equivalence class of $\psi_k \in S(U)$.

Lemma 4.2. *The collection of modules* $\mathrm{Sol}(U)$ *for* U *an open subset of* M, *together with the natural restriction maps* $r_{U,V} : \mathrm{Sol}(V) \to \mathrm{Sol}(U)$ *for* U, V *open subsets of* M *such that* $U \subset V$, *define a complete presheaf.*

Thus, we obtain a sheaf Sol over M, called the sheaf of microlocal solutions on M.

4B. *The sheaf of microlocal solutions.* One can show that if the principal symbol f_0 of T_k does not vanish on U, then $\mathrm{Sol}(U) = \{0\}$. Equivalently, if $\psi_k \in \mathscr{H}_k$ satisfies $T_k \psi_k = 0$, then its microsupport is contained in the level $\Gamma_0 = f_0^{-1}(0)$. This implies the following lemma.

Lemma 4.3. *Let* Ω *be an open subset of* Γ_0; *write* $\Omega = U \cap \Gamma_0$, *where* U *is an open subset of* M. *Then the restriction map*

$$r_\Omega : \mathrm{Sol}(U) \to \mathfrak{F}_U(\Omega) := r_\Omega(\mathrm{Sol}(U)), \quad [\psi_k] \mapsto [\psi_k|_\Omega]$$

is an isomorphism of $\mathbb{C}_k$-*modules.*

We want to define a new sheaf $\mathfrak{F} \to \Gamma_0$ that still describes the microlocal solutions of (3). In order to do so, we will check that the module $\mathfrak{F}_U(\Omega)$ does not depend on the open set U such that $\Omega = \Gamma_0 \cap U$. We first prove:

Lemma 4.4. *Let $U, \widetilde{U}$ be two open subsets of M such that $\Omega = U \cap \Gamma_0 = \widetilde{U} \cap \Gamma_0$. Then there exists an isomorphism between $\mathrm{Sol}(U)$ and $\mathrm{Sol}(\widetilde{U})$ commuting with the restriction maps.*

Proof. Assume that U and $\widetilde{U}$ are distinct and set $V = U \cap \widetilde{U}$; of course $\Omega \subset V$. Write $\widetilde{U} = V \cup W$ where the open set W is such that there exists an open set $X \subset V$ containing Ω such that $W \cap X = \varnothing$. Let χ_V, χ_W be a partition of unity subordinate to $\widetilde{U} = V \cup W$; in particular, $\chi_V(x) = 1$ whenever $x \in X$. One can show that the class $F_{\chi_V}(\psi_k) = [\chi_V \psi_k]$ belongs to $\mathrm{Sol}(\widetilde{U})$. We claim that the map F_{χ_V} is an isomorphism with the required property. $\qquad\square$

From these two lemmas, we deduce:

Proposition 4.5. *Let $U, \widetilde{U}$ be two open subsets of M such that $\Omega = U \cap \Gamma_0 = \widetilde{U} \cap \Gamma_0$. Then $\mathfrak{F}_U(\Omega) = \mathfrak{F}_{\widetilde{U}}(\Omega)$.*

This allows to define a sheaf $\mathfrak{F} \to \Gamma_0$, which will be called the sheaf of microlocal solutions over Γ_0. Let us point out that so far, we have made no assumption on the structure (regularity) of the level Γ_0.

4C. *Regular case.* Consider a point $m \in \Gamma_0$ which is regular for the principal symbol f_0. Then there exists a symplectomorphism χ between a neighborhood of m in M and a neighborhood of the origin in $\mathbb{R}^2$ such that $(f_0 \circ \chi^{-1})(x, \xi) = \xi$. We can quantize this symplectomorphism by means of a Fourier integral operator [Boutet de Monvel and Guillemin 1981; Zelditch 1997; Charles 2003b; Le Floch 2014]: there exists an admissible sequence of operators $U_k^{(m)} : \mathscr{C}^\infty(\mathbb{R}^2, L_0^k) \to \mathscr{C}^\infty(M, L^k \otimes K)$ such that

$$U_k^{(m)}(U_k^{(m)})^* \sim \Pi_k \quad \text{near } m$$

and

$$(U_k^{(m)})^* U_k^{(m)} \sim \Pi_k^0, \quad (U_k^{(m)})^* T_k U_k^{(m)} \sim S_k \quad \text{near } 0,$$

where S_k is the Toeplitz operator

$$S_k = \frac{i}{\sqrt{2}}\left(z - \frac{1}{k}\frac{d}{dz}\right),$$

which means that $S_k u = \frac{i}{\sqrt{2}}\left(zf - \frac{1}{k}\frac{df}{dz}\right)\psi^k$ if $u = f\psi^k$. Consider the element Φ_k of $\mathscr{C}^\infty(\mathbb{R}^2, L_0^k)$ given by

$$\Phi_k(z) = \exp(kz^2/2)\psi^k(z), \quad \psi(z) = \exp\left(-\tfrac{1}{2}|z|^2\right);$$

it satisfies $S_k \Phi_k = 0$. Choosing a suitable cutoff function η and setting $\Phi_k^{(m)} = \Pi_k^0(\eta \Phi_k)$, we obtain an admissible sequence $\Phi_k^{(m)}$ of elements of $\mathscr{B}_k$ microlocally equal to Φ_k near the origin and generating the $\mathbb{C}_k$-module of microlocal solutions of $S_k u_k = 0$ near the origin.

Proposition 4.6. *The $\mathbb{C}_k$-module of microlocal solutions of (3) near m is free of rank 1,[2] generated by $U_k^{(m)}\Phi_k^{(m)}$.*

---

[2]We recall that this means that this module admits a basis with one element.

This is a slightly modified version of Proposition 3.6 of [Charles 2003b], in which the normal form is achieved on the torus instead of the complex plane.

Thus, if Γ_0 contains only regular points of the principal symbol f_0, then $\mathfrak{F} \to \Gamma_0$ is a sheaf of free $\mathbb{C}_k$-modules of rank 1; in particular, this implies that $\mathfrak{F} \to \Gamma_0$ is a flat sheaf, thus characterized by its Čech holonomy $\mathrm{hol}_{\mathfrak{F}}$.

4D. *Lagrangian sections.* In order to compute the holonomy $\mathrm{hol}_{\mathfrak{F}}$, we have to understand the structure of the microlocal solutions. For this purpose, a family of solutions of particular interest is given by Lagrangian sections; let us define these. Consider a curve $\Gamma \subset \Gamma_0$ containing only regular points, and let $j : \Gamma \to M$ be the embedding of Γ into M. Let U be an open set of M such that $U_\Gamma = j^{-1}(U \cap \Gamma)$ is contractible; then there exists a flat unitary section t_Γ of $j^*L \to U_\Gamma$. Now, consider a formal series

$$\sum_{\ell \geq 0} \hbar^\ell g_\ell \in \mathscr{C}^\infty(U_\Gamma, j^*K)[\![\hbar]\!].$$

Let V be an open subset of M such that $\overline{V} \subset U$. Then a sequence $\Psi_k \in \mathscr{H}_k$ is a *Lagrangian section* associated to (Γ, t_Γ) with symbol $\sum_{\ell \geq 0} \hbar^\ell g_\ell$ if

$$\Psi_k(m) = \left(\frac{k}{2\pi}\right)^{1/4} F^k(m)\tilde{g}(m, k) \quad \text{over } V,$$

where

- F is a section of $L \to U$ such that

$$j^*F = t_\Gamma \quad \text{and} \quad \bar{\partial}F = 0$$

  modulo a section vanishing to every order along $j(\Gamma)$, and $|F(m)| < 1$ if $m \notin j(\Gamma)$;

- $\tilde{g}(\,\cdot\,, k)$ is a sequence in $\mathscr{C}^\infty(U, K)$ admitting an asymptotic expansion $\sum_{\ell \geq 0} k^{-\ell} \tilde{g}_\ell$ in the $\mathscr{C}^\infty$ topology such that

$$j^*\tilde{g}_\ell = g_\ell \quad \text{and} \quad \bar{\partial}\tilde{g}_\ell = 0$$

  modulo a section vanishing at every order along $j(\Gamma)$.

Assume furthermore that Ψ_k is admissible in the sense that $\Psi_k(m)$ is uniformly $O(k^N)$ for some N and the same holds for its successive covariant derivatives. It is possible to construct such a section with given symbol $\sum_{\ell \geq 0} \hbar^\ell g_\ell$ (see [Charles 2006, §3]). Furthermore, if Ψ_k is a nonzero Lagrangian section, then the constants $c_k \in \mathbb{C}_k$ such that $c_k \Psi_k$ is still a Lagrangian section are the elements of the form

$$c_k = \rho(k) \exp(ik\phi(k)) + O(k^{-\infty}), \tag{4}$$

where $\rho(k), \phi(k) \in \mathbb{R}$ admit asymptotic expansions of the form $\rho(k) = \sum_{\ell \geq 0} k^{-\ell} \rho_\ell$, $\phi(k) = \sum_{\ell \geq 0} k^{-\ell} \phi_\ell$.

Lagrangian sections are important because they provide a way to construct microlocal solutions. Indeed, if Ψ_k is a Lagrangian section over V associated to (Γ, t_Γ) with symbol $\sum_{\ell \geq 0} \hbar^\ell g_\ell$, then $T_k \Psi_k$ is also a Lagrangian section over V associated to (Γ, t_Γ), and one can in principle compute the elements $\hat{g}_\ell$, $\ell \geq 0$ of the formal expansion of its symbol as a function of the g_ℓ, $\ell \geq 0$ (by means of a stationary phase

expansion). This allows to solve (3) by prescribing the symbol of Ψ_k so that for every $\ell \geq 0$, $\hat{g}_\ell$ vanishes. Let us detail this for the two first terms.

Introduce a *half-form bundle* (δ, φ), that is, a line bundle $\delta \to M$ together with an isomorphism of line bundles $\varphi : \delta^{\otimes 2} \to \Lambda^{1,0} T^* M$. Since the first Chern class of M, which is equal to its Euler characteristic, is even, such a pair exists. Introduce the Hermitian holomorphic line bundle L_1 such that $K = L_1 \otimes \delta$. Define the *subprincipal form* κ as the 1-form on Γ such that

$$\kappa(X_{f_0}) = -f_1,$$

where X_{f_0} stands for the Hamiltonian vector field associated to f_0. Introduce the connection ∇^1 on $j^* L_1 \to \Gamma$ defined by

$$\nabla^1 = \nabla^{j^* L_1} + \frac{1}{i}\kappa,$$

with $\nabla^{j^* L_1}$ the connection induced by the Chern connection of L_1 on $j^* L_1$. Let δ_Γ be the restriction of δ to Γ; the map

$$\varphi_\Gamma : \delta_\Gamma^{\otimes 2} \to T^* \Gamma \otimes \mathbb{C}, \quad u \mapsto j^* \varphi(u)$$

is an isomorphism of line bundles. Define a connection ∇^{δ_Γ} on δ_Γ by

$$\nabla^{\delta_\Gamma}_X \sigma = \mathscr{L}^{\delta_\Gamma}_X \sigma,$$

where $\mathscr{L}^{\delta_\Gamma}_X$ is the first-order differential operator acting on sections of δ_Γ such that

$$\varphi_\Gamma\left(\mathscr{L}^{\delta_\Gamma}_X g \otimes g\right) = \tfrac{1}{2}\mathscr{L}_X \varphi_\Gamma(g^{\otimes 2})$$

for every section g; here, $\mathscr{L}$ stands for the standard Lie derivative of forms.

It was proved in [Charles 2006, Theorems 3.3 and 3.4] that $T_k \Psi_k$ is a Lagrangian section over V associated to t_Γ with symbol $(j^* f_0) g_0 + O(\hbar) = O(\hbar)$ (so Ψ_k satisfies (3) up to order $O(k^{-1})$) and that the subprincipal symbol of $T_k \Psi_k$ is

$$(j^* f_1) g_0 + \frac{1}{i}\left(\nabla^{j^* L_1}_{X_{f_0}} \otimes \mathrm{Id} + \mathrm{Id} \otimes \mathscr{L}^{\delta_\Gamma}_{X_{f_0}}\right) g_0.$$

Consequently, (3) is satisfied by Ψ_k up to order $O(k^{-2})$ if and only if

$$\left(f_1 + \frac{1}{i}\left(\nabla^{j^* L_1}_{X_{f_0}} \otimes \mathrm{Id} + \mathrm{Id} \otimes \mathscr{L}^{\delta_\Gamma}_{X_{f_0}}\right)\right) g_0 = 0 \quad \text{over } V \cap \Gamma. \tag{5}$$

This can be interpreted as a parallel transport equation: if we endow $j^* L_1 \otimes \delta_\Gamma$ with the connection induced from ∇^1 and ∇^{δ_Γ}, (5) means that g_0 is flat.

4E. *Holonomy.* We now assume that Γ_0 is connected (otherwise, one can consider connected components of Γ_0) and contains only regular points; it is then a smooth closed curve embedded in M. We would like to compute the holonomy of the sheaf $\mathfrak{F} \to \Gamma_0$.

Proposition 4.7. *The holonomy* $\mathrm{hol}_{\mathfrak{F}}(\Gamma_0)$ *is of the form*

$$\mathrm{hol}_{\mathfrak{F}}(\Gamma_0) = \exp(ik\Theta(k)) + O(k^{-\infty}), \tag{6}$$

where $\Theta(k)$ is real-valued and admits an asymptotic expansion of the form $\Theta(k) = \sum_{\ell \geq 0} k^{-\ell} \Theta_\ell$.

In particular, this means that if we consider another set of solutions to compute the holonomy, we only have to keep track of the phases of the transition constants.

Proof. Cover Γ_0 by a finite number of open subsets Ω_α in which the normal form introduced before Proposition 4.6 applies, and let U_k^α and Φ_k^α be as in this proposition. We obtain a family u_k^α of microlocal solutions; observe that for each α, u_k^α is a Lagrangian section associated to Γ. Hence, if $\Omega_\alpha \cap \Omega_\beta$ is nonempty, the unique (modulo $O(k^{-\infty})$) constant $c_k^{\alpha\beta} \in \mathbb{C}_k$ such that $u_k^\alpha = c_k^{\alpha\beta} u_k^\beta$ on $\Omega_\alpha \cap \Omega_\beta$ is of the form given in (4):

$$c_k^{\alpha\beta} = \rho^{\alpha\beta}(k) \exp(ik\phi^{\alpha\beta}(k)) + O(k^{-\infty}).$$

But if m belongs to $\Omega_\alpha \cap \Omega_\beta$, then near m we have $u_k^\alpha \sim U_k^\alpha \Phi_k^{(m)}$ and $u_k^\beta \sim U_k^\beta \Phi_k^{(m)}$, where $\Phi_k^{(m)}$ is an admissible sequence of elements of $\mathcal{B}_k$ microlocally equal to Φ_k near the origin. Therefore, we have

$$c_k^{\alpha\beta} \Phi_k^{(m)} = (U_k^\beta)^{-1} U_k^\alpha \Phi_k^{(m)} + O(k^{-\infty}),$$

and the fact that the operators U_k^α, U_k^β are microlocally unitary yields $|c_k^{\alpha\beta}|^2 = 1 + O(k^{-\infty})$. This implies that the coefficients $\rho_\ell^{\alpha\beta}$ in the asymptotic expansion of $\rho^{\alpha\beta}(k)$ vanish for $\ell \geq 1$, which gives the result. $\square$

Let us be more specific and compute the first terms of this asymptotic expansion. Consider a finite cover $(\Omega_\alpha)_\alpha$ of Γ_0 by open subsets with $j^{-1}(\Omega_\alpha)$ contractible, and endow a neighborhood of each Ω_α in M with a nontrivial microlocal solution Ψ_k^α which is a Lagrangian section. Choose a flat unitary section t_α of the line bundle $j^*L \to j^{-1}(\Omega_\alpha)$ and write, for $m \in \Omega_\alpha$,

$$\Psi_k^\alpha(m) = \left(\frac{k}{2\pi}\right)^{1/4} g_\alpha(m, k) t_\alpha^k(m),$$

where the section $g_\alpha(\cdot, k)$ of $j^*K \to \Omega_\alpha$ is the symbol of Ψ_k^α, whose principal symbol will be denoted by $g_\alpha^{(0)}$. Now, assume that $\Omega_\alpha \cap \Omega_\beta \neq \varnothing$; there exists a unique (up to $O(k^{-\infty})$) $c_k^{\alpha\beta} \in \mathbb{C}_k$ such that $\Psi_k^\alpha \sim c_k^{\alpha\beta} \Psi_k^\beta$ on $\Omega_\alpha \cap \Omega_\beta$.

Definition 4.8. Let $A, B \in M$ and γ be a piecewise smooth curve joining A and B; denote by $P_{A,B,\gamma} : L_A \to L_B$ the linear isomorphism given by parallel transport from A to B along γ. Given two sections s, t of $L \to M$ such that $s(A) \neq 0$ and $t(B) \neq 0$, define the *phase difference* between $s(A)$ and $t(B)$ along γ as the number

$$(\Phi_s(A) - \Phi_t(B))_\gamma = \arg \lambda_{A,B,\gamma} - c_0([A, B]) \in \mathbb{R}/2\pi\mathbb{Z},$$

where $\lambda_{A,B,\gamma}$ is the unique complex number such that $P_{A,B,\gamma}(s(A)) = \lambda_{A,B,\gamma} t(B)$ and $c_0([A, B])$ is the (phase of the) holonomy of γ in (L, ∇) (computed with respect to some fixed trivializations at A, B). Define in the same way the phase difference for two sections of $K \to M$, this time using the Chern connection of K.

Now, consider three points $A, B, C \in M$, and let γ_1 and γ_2 be piecewise smooth curves joining A to B and B to C, respectively. Let γ be the concatenation of γ_1 and γ_2. It is easily checked that

$$(\Phi_s(A) - \Phi_t(B))_{\gamma_1} + (\Phi_t(B) - \Phi_u(C))_{\gamma_2} = (\Phi_s(A) - \Phi_u(C))_\gamma$$

for three sections s, t, u of L. Furthermore, if γ is a closed curve and A is a point on γ, then the phase difference between $s(A)$ and $s(A)$ along γ is

$$(\Phi_s(A) - \Phi_s(A))_\gamma = 0$$

by definition of the holonomy c_0. This is why we write this number as a difference. Note that this still holds true if we change the set of trivializations used to compute c_0.

Coming back to our problem, denote by $\Phi_\alpha^{(-1)}(A) - \Phi_\beta^{(-1)}(B)$ the phase difference between $t_\alpha(A)$ and $t_\beta(B)$ along Γ_0 in L, and by $\Phi_\alpha^{(0)}(A) - \Phi_\beta^{(0)}(B)$ the phase difference between $g_\alpha^{(0)}(A)$ and $g_\beta^{(0)}(B)$ along Γ_0 in K. Let ζ be the path in Γ_0 starting at a point $A \in \Omega_\alpha$ and ending at $B \in \Omega_\alpha \cap \Omega_\beta$. Since t_α is flat and the principal symbol g_0 of Ψ_k^α satisfies (5), we have

$$\arg c_k^{\alpha\beta} = k\big(c_0(\zeta) + \Phi_\alpha^{(-1)}(A) - \Phi_\beta^{(-1)}(B)\big) + c_1(\zeta) + \mathrm{hol}_{\delta_{\Gamma_0}}(\zeta) + \Phi_\alpha^{(0)}(A) - \Phi_\beta^{(0)}(B) + O(k^{-1}),$$

where $c_1(\zeta)$ is the holonomy of ζ in (L_1, ∇^1) and $\mathrm{hol}_{\delta_{\Gamma_0}}(\zeta)$ is the holonomy of ζ in $(\delta_{\Gamma_0}, \nabla^{\delta_{\Gamma_0}})$ (both computed with respect to some fixed trivializations of L_1 and δ_{Γ_0} at A, B).

Thanks to the discussion above, we know that the term

$$k\big(\Phi_\alpha^{(-1)}(A) - \Phi_\beta^{(-1)}(B)\big) + \Phi_\alpha^{(0)}(A) - \Phi_\beta^{(0)}(B)$$

is a Čech coboundary. The values $c_0(\Gamma_0)$, $c_1(\Gamma_0)$ and $\mathrm{hol}_{\delta_{\Gamma_0}}(\Gamma_0)$ do not depend on the trivializations chosen for the computations. Moreover, one can check that $\nabla^{\delta_{\Gamma_0}}$ has holonomy in $\mathbb{Z}/2\mathbb{Z}$, represented by $\epsilon(\Gamma_0) \in \{0, 1\}$. Thus, we obtain:

Proposition 4.9. *The first two terms of the asymptotic expansion of the quantity $\Theta(k)$ defined in Proposition 4.7 are given by*

$$\Theta_0 = c_0(\Gamma_0)$$

and

$$\Theta_1 = c_1(\Gamma_0) + \epsilon(\Gamma_0)\pi.$$

Since one can construct a nontrivial microlocal solution over Γ_0 if and only if $\Theta(k) \in 2\pi\mathbb{Z}$, we recover the usual Bohr–Sommerfeld conditions.

Let us give another interpretation of the index ϵ. Consider a smooth closed curve γ immersed in M. Denote by $\iota : \gamma \to M$ this immersion, and by $\delta_\gamma = \iota^*\delta$ the pullback bundle over γ. Let $\tilde\iota : \delta_\gamma \to \delta$ be the natural lift of ι, and define $\tilde\iota^2 : \delta_\gamma^{\otimes 2} \to \delta^{\otimes 2}$ by the formula $\tilde\iota^2(u \otimes v) = \tilde\iota(u) \otimes \tilde\iota(v)$. The map

$$\varphi_\gamma : \delta_\gamma^{\otimes 2} \to T^*\gamma \otimes \mathbb{C}, \quad u \mapsto \iota^*\varphi(\tilde\iota^2(u))$$

is an isomorphism of line bundles. The set

$$\{u \in \delta_\gamma \mid \varphi_\gamma(u^{\otimes 2}) > 0\}$$

has one or two connected components. In the first case, we set $\epsilon(\gamma) = 1$, and in the second case $\epsilon(\gamma) = 0$. One can check that this definition coincides with the one above when γ is a smooth embedded closed curve. Notice that the value of $\epsilon(\gamma)$ only depends on the isotopy class of γ in M.

4F. *Spectral parameter dependence.* For spectral analysis, one has to do the same study as above, replacing the operator T_k with $T_k - \lambda$, $\lambda \in \mathbb{R}$; then it is natural to ask if the previous study can be done taking into account the dependence of the operator on the spectral parameter λ.

Assume that there exists a tubular neighborhood Ω of Γ such that for λ close enough to 0, the intersection $\Gamma_\lambda \cap \Omega$ is regular. Then we can construct microlocal solutions of $(T_k - \lambda)u_k = 0$ as Lagrangian sections depending smoothly on a parameter (see [Charles 2003b, §2.6]); these solutions are uniform in λ. We can then define all the previous objects with smooth dependence in λ. Proceeding this way, we obtain the parameter-dependent Bohr–Sommerfeld conditions, which we describe below.

Let I be an interval of regular values of the principal symbol f_0 of the operator. For $\lambda \in I$, denote by $\mathscr{C}_j(\lambda)$, $1 \le j \le N$, the connected components of $f_0^{-1}(\lambda)$ in such a way that for j fixed and $\lambda_1 \ne \lambda_2 \in I$, $\mathscr{C}_j(\lambda_1)$ and $\mathscr{C}_j(\lambda_2)$ belong to the same connected component of $f_0^{-1}(I)$. Observe that $\mathscr{C}_j(\lambda)$ is a smooth embedded closed curve, endowed with the orientation depending continuously on λ given by the Hamiltonian flow of f_0. Define the *principal action* $c_0^{(j)} \in \mathscr{C}^\infty(I)$ in such a way that the parallel transport in L along $\mathscr{C}_j(\lambda)$ is the multiplication by $\exp(ic_0^{(j)}(\lambda))$. Define the *subprincipal action* $c_1^{(j)}$ in the same way, replacing L by L_1 and using the connection ∇^1 (depending on λ) described above. Finally, set $\epsilon_\lambda^{(j)} = \epsilon(\mathscr{C}_j(\lambda))$; in fact, $\epsilon_\lambda^{(j)}$ is a constant $\epsilon_\lambda^{(j)} = \epsilon^{(j)}$ for λ in I. Fix E in I; the Bohr–Sommerfeld conditions (see [Charles 2006] for more details) state that there exists $\eta > 0$ such that the intersection of the spectrum of T_k with $[E - \eta, E + \eta]$ modulo $O(k^{-\infty})$ is the union of the spectra σ_j, $1 \le j \le N$, where the elements of σ_j are the solutions of

$$g^{(j)}(\lambda, k) \in 2\pi k^{-1}\mathbb{Z},$$

where $g^{(j)}(\,\cdot\,, k)$ is a sequence of functions of $\mathscr{C}^\infty(I)$ admitting an asymptotic expansion

$$g^{(j)}(\,\cdot\,, k) = \sum_{\ell \ge 0} k^{-\ell} g_\ell^{(j)}$$

with coefficients $g_\ell^{(j)} \in \mathscr{C}^\infty(I)$. Furthermore, one has

$$g_0^{(j)}(\lambda) = c_0^{(j)}(\lambda) \quad \text{and} \quad g_1^{(j)}(\lambda) = c_1^{(j)}(\lambda) + \epsilon^{(j)}\pi.$$

5. Microlocal normal form

5A. *Normal form on the Bargmann side.* Let P_k be the operator defined by

$$P_k = \frac{i}{2}\left(z^2 - \frac{1}{k^2}\frac{\partial^2}{\partial z^2}\right)$$

with domain $\mathbb{C}[z] \subset \mathscr{B}_k$; it is a Toeplitz operator with normalized symbol $p_0(x, \xi) = x\xi$. We will use this operator to understand the behavior of A_k when acting on sections localized near each s_j, $1 \le j \le n$. In fact, we study the operator $A_k - E$, where $E \in \mathbb{R}$ is allowed to vary in a neighborhood of zero.

Let $j \in [\![1, n]\!]$. The isochore Morse lemma [Colin de Verdière and Vey 1979] yields a symplectomorphism χ_E from a neighborhood of s_j in M to a neighborhood of the origin in $\mathbb{R}^2$, depending smoothly on

E, and a smooth function g_j^E, again depending smoothly on E, such that

$$((a_0 - E) \circ \chi_E^{-1})(x, \xi) = g_j^E(x\xi)$$

and $(g_j^E)'(0) \neq 0$. Using a Taylor formula, one can write

$$g_j^E(t) = w_j^E(t)(t - f_j(E))$$

with w_j^E smooth, depending smoothly on E, and such that $w_j^E(0) \neq 0$, and with f_j a smooth function of E with $f_j(0) = 0$. This symplectic normal form can be quantized to the following semiclassical normal form.

Proposition 5.1. *Fix $j \in [\![1, n]\!]$. Then there exist a smooth function f_j, a Fourier integral operator $U_k^E : \mathscr{B}_k \to \mathscr{H}_k$, a Toeplitz operator W_k^E, elliptic at 0, and a sequence of smooth functions $\varepsilon_j(\cdot, k)$ admitting an asymptotic expansion $\varepsilon_j(E, k) = \sum_{\ell=0}^{+\infty} k^{-\ell} \varepsilon_j^{(\ell)}(E)$ such that*

$$(U_k^E)^*(A_k - E)U_k^E \sim W_k^E\left(P_k - f_j(E) - k^{-1}\varepsilon_j(E, k)\right)$$

microlocally near s_j. Furthermore,

- *U_k and W_k depend smoothly on E,*

- *$f_j(E)$ is the value of $x\xi$ whenever $(x, \xi) = \chi_E(m)$ for $m \in \Gamma_E$, and*

- *the first term of the asymptotic expansion of $\varepsilon_j(0, k)$ is given by*

$$\varepsilon_j^{(0)}(0) = \frac{-a_1(s_j)}{|\det(\mathrm{Hess}(a_0)(s_j))|^{1/2}},$$

where $\mathrm{Hess}(a_0)(s_j)$ is the Hessian of a_0 at s_j.

The proof is an adaptation of the one in [Colin de Verdière and Parisse 1994a, §3] to the Toeplitz setting; see also [Le Floch 2014, Theorem 5.3] for a similar result in the elliptic case.

5B. *Link with the pseudodifferential setting.* Now we use the Bargmann transform to understand the structure of the space of microlocal solutions of $P_k - E = 0$.

Lemma 5.2. *For $u \in \mathscr{S}(\mathbb{R})$, one has*

$$B_k^* P_k B_k u = \frac{1}{ik}(x\partial_x + 1)u.$$

From now on, we will denote by S_k the pseudodifferential operator $(1/ik)(x\partial_x + 1)$. This correspondence will allow us to understand the space of microlocal solutions of $P_k - E$ on a neighborhood of the origin. Let us recall the results of [Colin de Verdière and Parisse 1994a; 1994b] that will be useful for our study.

Proposition 5.3 [Colin de Verdière and Parisse 1994a, Proposition 3]. *Let E be such that $|E| < 1$. The space of microlocal solutions of $(S_k - E)u_k = 0$ on $Q = [-1, 1]^2$ is a free $\mathbb{C}_k$-module of rank 2.*

Moreover, we know two bases of this module. Indeed, let $\mathscr{F}_k$ be the semiclassical Fourier transform:

$$(\mathscr{F}_k u)(\xi) = \frac{k}{2\pi} \int_{\mathbb{R}} \exp(-ikx\xi) u(x)\, dx;$$

then the tempered distributions $v_{k,E}^{(j)}$, $j \in [\![1, 4]\!]$, defined as

$$v_{k,E}^{(1),(2)}(x) = \mathbf{1}_{\mathbb{R}_{\pm}^*}(x) \exp\big(\big(-\tfrac{1}{2} + ikE\big) \ln |x|\big),$$

$$v_{k,E}^{(3),(4)}(x) = \mathscr{F}_k^{-1}\Big(\mathbf{1}_{\mathbb{R}_{\pm}^*}(\xi) \exp\big(\big(-\tfrac{1}{2} + ikE\big) \ln |\xi|\big)\Big)(x),$$

are exact solutions of the equation $(S_k - E)v_{k,E}^{(j)} = 0$; better than that, the pairs $(v_{k,E}^{(1)}, v_{k,E}^{(2)})$ and $(v_{k,E}^{(3)}, v_{k,E}^{(4)})$ each form a basis of the space of solutions of this equation. Now, choose a compactly supported function $\chi \in \mathscr{C}^\infty(\mathbb{R})$ with constant value 1 on a neighborhood of $I = [-1, 1]$ and vanishing outside $2I$. Define the pseudodifferential operator Π_Q by

$$\Pi_Q u(x) = \frac{k}{2\pi} \int_{\mathbb{R}^2} \exp(ik(x-y)\xi) \chi(\xi) \chi(y) u(y)\, dy\, d\xi.$$

Then Π_Q maps $\mathscr{D}'(\mathbb{R})$ into $\mathscr{S}(\mathbb{R})$, and $\Pi_Q \sim \mathrm{Id}$ on Q. Set

$$w_{k,E}^{(j)} = \Pi_Q v_k^{(j)};$$

then the $w_{k,E}^{(j)}$, $j \in [\![1, 4]\!]$, belong to $\mathscr{S}(\mathbb{R})$, and are microlocal solutions of $(S_k - E)w_{k,E}^{(j)} = 0$ on Q. The matrix of the change of basis from $(w_{k,E}^{(3)}, w_{k,E}^{(4)})|_Q$ to $(w_{k,E}^{(1)}, w_{k,E}^{(2)})|_Q$ is given by

$$M_k(E) = \mu_k(E) \begin{pmatrix} 1 & i\exp(-\pi kE) \\ i\exp(-\pi kE) & 1 \end{pmatrix} + O(k^{-\infty}), \tag{7}$$

with

$$\mu_k(E) = \frac{1}{\sqrt{2\pi}} \Gamma\big(\tfrac{1}{2} + ikE\big) \exp\Big(\frac{\pi}{4}(2kE - i) - ikE \ln k\Big).$$

5C. *Microlocal solutions of $(P_k - E)u_k = 0$.* Now, consider the Bargmann transforms of the sequences $w_{k,E}^{(j)}$: $u_{k,E}^{(j)} = B_k w_{k,E}^{(j)}$. Propositions 5.3 and 3.5 yield:

Proposition 5.4. *For E such that $|E| < 1$, the space of microlocal solutions of $(P_k - E)u_k = 0$ on $Q = [-1, 1]^2 \subset \mathbb{C}$ is a free $\mathbb{C}_k$-module of rank 2. Moreover, the pairs $(u_{k,E}^{(1)}, u_{k,E}^{(2)})$ and $(u_{k,E}^{(3)}, u_{k,E}^{(4)})$ are two bases of this module; the transfer matrix is given by* (7).

Remark. The sections $u_{k,E}^{(j)}$, $j = 1, \ldots, 4$, can be written in terms of parabolic cylinder functions. Nonnenmacher and Voros [1997] studied these functions in order to understand the behavior of the generalized eigenfunctions of P_k; the result of this subtle analysis, based on Stokes lines techniques, was not exactly what we needed here, and this is partly why we chose to use the microlocal properties of the Bargmann transform instead.

6. Bohr–Sommerfeld conditions

To obtain the Bohr–Sommerfeld conditions, we will recall the reasoning of [Colin de Verdière and Parisse 1999], and will also refer to [Colin de Verdière and Vũ Ngọc 2003]. Since the general approach is the same, we only recall the main ideas and focus on what differs in the Toeplitz setting.

6A. *The sheaf of standard bases.* As in Section 4, introduce the sheaf $(\mathfrak{F}, \Gamma_0)$ of microlocal solutions of $A_k \psi_k = 0$ over Γ_0; we recall that a global nontrivial microlocal solution corresponds to a global nontrivial section of this sheaf. However, since the topology of Γ_0 is much more complicated than in the regular case, the condition for the existence of such a section is not as simple as saying that a holonomy must be trivial. In particular, we have to handle what happens at critical points. To overcome this difficulty, the idea is to introduce a new sheaf over Γ_0 that will contain all the information we need to construct a global nontrivial microlocal solution; roughly speaking, this new sheaf can be thought of as the limit of the sheaf $\mathfrak{F} \to \Gamma_E$ of microlocal solutions over regular levels as E goes to 0.

Following [Colin de Verdière and Parisse 1999], we introduce a sheaf $(\mathfrak{L}, \Gamma_0)$ of free $\mathbb{C}_k$-modules of rank 1 over Γ_0 as follows: to each point $m \in \Gamma_0$, associate the free module $\mathfrak{L}(m)$ generated by *standard bases* at m. If m is a regular point, a standard basis is any basis of the space of microlocal solutions near m. At a critical point s_j, we define a standard basis in the following way. The $\mathbb{C}_k$-module of microlocal solutions near s_j is free of rank 2; moreover, it is the graph of a linear function. Indeed, number the four local edges near s_j with cyclic order 1, 3, 2, 4, so that the edges e_1, e_2 are the ones that leave s_j. Let us denote by $\mathrm{Sol}(e_1 e_2)$ and $\mathrm{Sol}(e_3 e_4)$ the modules of microlocal solutions over the disjoint union of the local unstable edges e_1, e_2 and stable edges e_3, e_4, respectively. $\mathrm{Sol}(e_1 e_2)$ and $\mathrm{Sol}(e_3 e_4)$ are free modules of rank 2, and there exists a linear map $T_j : \mathrm{Sol}(e_3 e_4) \to \mathrm{Sol}(e_1 e_2)$ such that u is a solution near s_j if and only if its restrictions satisfy $u|_{\mathrm{Sol}(e_1 e_2)} = T_j u|_{\mathrm{Sol}(e_3 e_4)}$. Equivalently, given two solutions on the entering edges, there is a unique way to obtain two solutions on the leaving edges by passing the singularity. One can choose a basis element for each $\mathfrak{F}(e_i)$, $i \in [\![1, 4]\!]$, and express T_j as a 2×2 matrix (defined modulo $O(k^{-\infty})$); one can show that the entries of this matrix are all nonvanishing. An argument of elementary linear algebra shows that, once the matrix T_j is chosen, the basis elements of the modules $\mathfrak{F}(e_i)$ are fixed up to multiplication by the same factor; this means that for T_j fixed, the $\mathbb{C}_k$-module of basis elements is of rank 1. Moreover, the study of the previous section implies that there exists a choice of basis elements such that T_j has the following expression:

$$T_j = \exp\left(-\frac{i\pi}{4}\right) \mathscr{E}_k(\varepsilon_j(0, k)) \begin{pmatrix} 1 & i \exp(-\pi \varepsilon_j(0, k)) \\ i \exp(-\pi \varepsilon_j(0, k)) & 1 \end{pmatrix}, \tag{8}$$

where

$$\mathscr{E}_k(t) = \frac{1}{\sqrt{2\pi}} \Gamma\left(\tfrac{1}{2} + it\right) \exp\left(t\left(\frac{\pi}{2} - i \ln k\right)\right). \tag{9}$$

This allows us to call the choice of the basis elements of $\mathfrak{F}(e_i)$ a standard basis if T_j is given by (8).

$(\mathcal{L}, \Gamma_0)$ is a locally free sheaf of rank-1 $\mathbb{C}_k$-modules, and its transition functions are constants. Hence, it is flat, thus characterized by its holonomy

$$\mathrm{hol}_{\mathcal{L}} : \mathrm{H}_1(\Gamma_0) \to \mathbb{C}_k.$$

In terms of Čech cohomology, if γ is a cycle in Γ_0 and $\Omega_1, \ldots, \Omega_\ell$ is an ordered sequence of open sets covering the image of γ, each Ω_i being equipped with a standard basis u_i (at a critical point, we make the abusive correspondence between a standard basis and its elements), then

$$\mathrm{hol}_{\mathcal{L}}(\gamma) = x_{1,2} \cdots x_{\ell-1,\ell} x_{\ell,1}, \tag{10}$$

where $x_{i,j} \in \mathbb{C}_k$ is such that $u_i = x_{i,j} u_j$ on $\Omega_i \cap \Omega_j$.

Now, cut $n+1$ edges of Γ_0, each one corresponding to a cycle γ_i in a basis $(\gamma_1, \ldots, \gamma_{n+1})$ of $\mathrm{H}_1(\Gamma_0, \mathbb{Z})$, in such a way that the remaining graph is a tree $\mathcal{T}$. Then the sheaf $(\mathcal{L}, \mathcal{T})$ has a nontrivial global section. The conditions to obtain a nontrivial global section of the sheaf $(\mathfrak{F}, \Gamma_0)$ of microlocal solutions on Γ_0 are given in the following theorem. They were already present in the work of Colin de Verdière and Parisse in the case of pseudodifferential operators, but the fact that they extend to our setting is a consequence of the results obtained in the previous sections.

Theorem 6.1. *The sheaf $(\mathfrak{F}, \Gamma_0)$ has a nontrivial global section if and only if the following linear system of $3n + 1$ equations with $3n + 1$ unknowns $(x_\alpha \in \mathbb{C}_k)_{\alpha \in \{\text{edges of } \mathcal{T}\}}$ has a nontrivial solution:*

(1) *If the edges $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ connect at s_j (with the same convention as before for the labeling of the edges), then*

$$\begin{pmatrix} x_{\alpha_3} \\ x_{\alpha_4} \end{pmatrix} = T_j \begin{pmatrix} x_{\alpha_1} \\ x_{\alpha_2} \end{pmatrix}$$

(2) *If α and β are the extremities of a cut cycle γ_i, then*

$$x_\alpha = \mathrm{hol}_{\mathcal{L}}(\gamma_i) x_\beta,$$

where the following orientation is assumed: γ_i can be represented as a closed path starting on the edge α and ending on the edge β.

Proof. It follows from Propositions 4.6 and 5.4 that the proof can be directly adapted from the one of [Colin de Verdière and Vũ Ngọc 2003, Theorem 2.7]. □

6B. *Singular invariants.* Of course, in order to use this result, it remains to compute the holonomy $\mathrm{hol}_{\mathcal{L}}$. For this purpose, let us introduce some geometric quantities close to the ones used to express the regular Bohr–Sommerfeld conditions. Let γ be a cycle in Γ_0, and denote by $s_{j_m}, m = 1, \ldots, p$, the critical points contained in γ.

Definition 6.2 (singular subprincipal action). Decompose γ as a concatenation of smooth paths and paths containing exactly one critical point; if A and B are the ordered endpoints of a path, we will call it $[A, B]$. Define the subprincipal action $\tilde{c}_1(\gamma)$ as the sum of the contributions of these paths, given by the following rules:
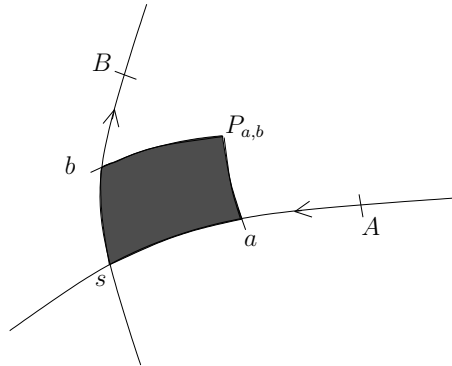
Figure 1. Computation of $\tilde{c}_1([A, B])$.

• If $[A, B]$ contains only regular points, its contribution to the singular subprincipal action is

$$\tilde{c}_1([A, B]) = c_1([A, B]),$$

as in the regular case (see Section 4E for the definition of $c_1([A, B])$);

• If $[A, B]$ contains the singular point s and is smooth at s, then

$$\tilde{c}_1([A, B]) = \lim_{a,b \to s} (c_1([A, a]) + c_1([b, B])),$$

where a and b lie on the same branches as A and B, respectively;

• If $[A, B]$ contains the singular point s and is not smooth at s, we set

$$\tilde{c}_1([A, B]) = \lim_{a,b \to s} \left(c_1([A, a]) + c_1([b, B]) \pm \varepsilon_s^{(0)} \ln \left| \int_{P_{a,b}} \omega \right| \right), \tag{11}$$

where $P_{a,b}$ is the parallelogram (defined in any coordinate system) built on the vectors $\overrightarrow{sa}$ and $\overrightarrow{sb}$, $\pm = +$ if $[A, B]$ is oriented according to the flow of X_{a_0}, $\pm = -$ otherwise, and

$$\varepsilon_s^{(0)} = \frac{-a_1(s)}{|\det(\mathrm{Hess}(a_0)(s))|^{1/2}},$$

as before.

Definition 6.3 (singular index). Let $(\gamma_t)_t$ be a continuous family of immersed closed curves such that $\gamma_0 = \gamma$ and γ_t is smooth for $t > 0$. Then the function $t \mapsto \epsilon(\gamma_t)$, $t > 0$, is constant; we denote by ϵ its value. We define the singular index $\tilde{\epsilon}(\gamma)$ by setting

$$\tilde{\epsilon}(\gamma) = \epsilon + \sum_{m=1}^{p} \frac{\rho_m}{4}, \tag{12}$$

where $\rho_m = 0$ if γ is smooth at s_{j_m}, $\rho_m = +1$ if at s_{j_m}, γ turns in the direct sense with respect to the cyclic order $(1, 3, 2, 4)$ of the local edges, and $\rho_m = -1$ otherwise.

Observe that both $\tilde{c}_1$ and $\tilde{\epsilon}$ define $\mathbb{Z}$-linear maps on $H_1(\Gamma_0, \mathbb{Z})$.

Theorem 6.4. *Let γ be a cycle in Γ_0. Then the holonomy $\mathrm{hol}_{\mathfrak{L}}(\gamma)$ of γ in $\mathfrak{L}$ has the form*

$$\mathrm{hol}_{\mathfrak{L}}(\gamma) = \exp(ik\theta(\gamma, k)), \tag{13}$$

where $\theta(\gamma, k)$ admits an asymptotic expansion in nonpositive powers of k. Moreover, if we denote by $\theta(\gamma, k) = \sum_{\ell \geq 0} k^{-\ell} \theta_\ell(\gamma)$ this expansion, the first two terms are given by the formulas

$$\theta_0(\gamma) = c_0(\gamma), \quad \theta_1(\gamma) = \tilde{c}_1(\gamma) + \tilde{\epsilon}(\gamma)\pi. \tag{14}$$

Proof. We just prove here that the holonomy has the claimed behavior. It is enough to show that one can choose a finite open cover $(\Omega_\alpha)_\alpha$ of γ and a section u_k^α of $\mathfrak{L} \to \Omega_\alpha$ for which the transition constants $c_k^{\alpha\beta}$ have the required form. On the edges of γ, this follows from the analysis of Section 4. At a vertex, we choose the standard basis $U_k^0 u_{k,\varepsilon_j(0,k)}^{(j)}$, where $u_{k,E}^{(j)}$ is defined in Section 5C and U_k^E is the operator of Proposition 5.1; to conclude, we observe that the restrictions of these sections to the corresponding edge are Lagrangian sections. $\square$

6C. *Computation of the singular holonomy.*

This section is devoted to the proof of the second part of Theorem 6.4. We use the method of [Colin de Verdière and Vũ Ngọc 2003], but of course, our case is simpler, because in the latter, the authors investigated the case of singularities in (real) dimension 4 (for pseudodifferential operators). Let us work on microlocal solutions of the equation

$$(A_k - E)u_k = 0, \tag{15}$$

where E varies in a small interval I containing the critical value 0. The critical value separates I into two open sets I^+ and I^-, with the convention $I^\pm = I \cap \mathbb{R}_\pm^*$. Let $D^\pm = I^\pm \cup \{0\}$, and let $\mathfrak{C}^\pm$ be the set of connected components of the open set $a_0^{-1}(I^\pm)$. The smooth family of circles in the component $p^\pm$ is denoted by $\mathscr{C}_{p^\pm}(E)$, $E \in I^\pm$.

As in Section 4, for $E \neq 0$, we denote by $(\mathfrak{F}, \Gamma_E)$ the sheaf of microlocal solutions of (15) on Γ_E; remember that it is a flat sheaf of rank-1 $\mathbb{C}_k$-modules, characterized by its Čech holonomy $\mathrm{hol}_{\mathfrak{F}}$. The idea is to let E go to 0 and compare this holonomy to the holonomy of the sheaf $\mathfrak{L} \to \Gamma_0$.

Definition 6.5. Near each critical point s_j, we consider two families of points $A_j(E)$ and $B_j(E)$ in $\mathscr{C}^\infty(D^\pm, \bar{p}^\pm \setminus \{s_j\})$ lying on $\mathscr{C}_{p^\pm}(E)$, and such that $A_j(0)$ and $B_j(0)$ lie in the stable and unstable manifolds, respectively. Endow a small neighborhood of A_j (resp. B_j) with a microlocal solution u_{A_j} (resp. u_{B_j}) of (15) which is a Lagrangian section uniform in $E \in D^\pm$. Define the quantity $\Theta([A_j(E), B_j(E)], k)$ as the phase of the Čech holonomy of the path $[A_j(E), B_j(E)] \subset \Gamma_E$ joining $A_j(E)$ and $B_j(E)$ in the sheaf $(\mathfrak{F}, \Gamma_E)$ computed with respect to u_{A_j} and u_{B_j}. Define in the same way the quantity $\Theta([B_j(E), A_{j'}(E)], k)$ for the path joining $B_j(E)$ and $A_{j'}(E)$.

Note that if we change the sections u_{A_j} and $u_{B_{j'}}$, the phase of the holonomy is modified by an additive term admitting an asymptotic expansion in $k\,\mathscr{C}^\infty(D^\pm)[\![k^{-1}]\!]$. The singular behavior of the holonomy is thus preserved; moreover, the added term is a Čech coboundary, and hence does not change the value of the holonomy along a closed path.
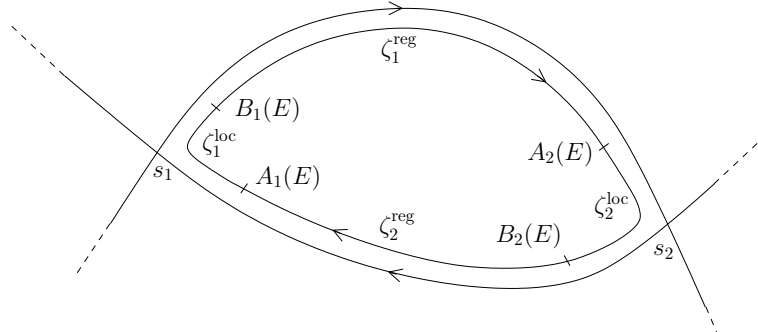
Figure 2. Regular and local paths.

Then, we consider continuous families of paths $(\zeta_E)_{E \in D^{\pm}}$ drawn on a circle $\mathscr{C}_{p^{\pm}}(E)$ and whose endpoints are some of the $A_j(E)$ and $B_{j'}(E)$ of the previous definition. We say that ζ_E is

- *regular* if ζ_0 does not contain any of the critical points s_j,
- *local* if ζ_0 contains exactly one critical point,

and we consider only these two types of paths. The following proposition implies that a path that is local in the above sense can always be assumed to be local in the sense that it is included in a small neighborhood of the critical point that it contains.

Proposition 6.6. *If* $\zeta_E = [B_j(E), A_{j'}(E)]$ *is a regular path, then the map* $E \mapsto \Theta(\zeta_E, k)$ *belongs to* $\mathscr{C}^{\infty}(D^{\pm})$ *and admits an asymptotic expansion in* $k\mathscr{C}^{\infty}(D^{\pm})[\![k^{-1}]\!]$. *This expansion starts as follows*:

$$\Theta(\zeta_E, k) = k\big(c_0(\zeta_E) + \Phi_{B_j(E)}^{(-1)}(B_j(E)) - \Phi_{A_{j'}(E)}^{(-1)}(A_{j'}(E))\big)$$
$$+ c_1(\zeta_E) + \mathrm{hol}_{\delta_{\zeta_E}}(\zeta_E) + \Phi_{B_j(E)}^{(0)}(B_j(E)) - \Phi_{A_{j'}(E)}^{(0)}(A_{j'}(E)) + O(k^{-1}); \quad (16)$$

see Section 4E for the notation.

In order to study the behavior of the holonomy of a local path with respect to E, we use the parameter-dependent normal form given by Proposition 5.1. Using the notation of this proposition, we will write $\mathfrak{e}_j(E, k) = f_j(E) + k^{-1}\varepsilon_j(E, k)$. Introduce the Bargmann transform $w_{k,E}^i$ of $v_{k,E}^i$, where

$$v_{k,E}^{1,2}(x) = \mathbf{1}_{\mathbb{R}_{\pm}^*}(x)|x|^{-1/2} \exp(ik\mathfrak{e}_j(E, k) \ln|x|),$$
$$v_{k,E}^{3,4}(x) = \mathscr{F}_k^{-1}\big(\mathbf{1}_{\mathbb{R}_{\pm}^*}(x)(\xi)|\xi|^{-1/2} \exp(ik\mathfrak{e}_j(E, k) \ln|\xi|)\big)(x).$$

Let $\widetilde{w}_{k,E}^i$ be a sequence having microsupport in a sufficiently small neighborhood of the origin and microlocally equal to $w_{k,E}^i$ on it; then $\widetilde{w}_{k,E}^i$ is a basis of the module of microlocal solutions of $P_k - \mathfrak{e}_j(E, k)$ near the image of the edge with label i by the symplectomorphism χ_E. Consequently, the section $\phi_{k,E}^{(i)} = U_k^E \widetilde{w}_{k,E}^i$, where U_k^E is the operator used for the normal form, is a basis of the module of microlocal solutions of (15) near the edge e_i. Moreover, it displays good behavior with respect to the spectral parameter.

Lemma 6.7. *The restriction of* $\phi_{k,E}^{(i)}$ *to a neighborhood of the edge numbered* i *is a Lagrangian section uniformly for* $E \in D^{\pm}$.

Proof. First, we prove using a parameter-dependent stationary phase lemma that $w_{k,E}^i$ is a Lagrangian section associated to the image of the i-th edge, uniformly in $E \in D^{\pm}$. We conclude by the fact that the image of a Lagrangian section depending smoothly on a parameter by a Fourier integral operator is a Lagrangian section depending smoothly on this parameter. □

We also recall the following useful lemma.

Lemma 6.8 [Colin de Verdière and Vũ Ngọc 2003, Lemma 2.18]. *Set* $\beta_j(E, k) = \frac{1}{2} + ik\mathfrak{e}_j(E, k)$ *and*

$$v_j^+ = \left(\frac{k}{2\pi}\right)^{1/2} \Gamma(\beta_j) \exp\left(-\beta_j \ln k - i\beta_j \frac{\pi}{2}\right),$$

$$v_j^- = \left(\frac{k}{2\pi}\right)^{1/2} \Gamma(\beta_j) \exp\left(-\beta_j \ln k + i\beta_j \frac{\pi}{2}\right),$$

so that

$$M_k(\mathfrak{e}_j(E, k)) = \begin{pmatrix} v_j^+(E, k) & v_j^-(E, k) \\ v_j^-(E, k) & v_j^+(E, k) \end{pmatrix}$$

where M_k *was defined in* (7). *Then, for any* $E \in I^{\pm}$,

$$-i \ln v_j^{\pm} = k\left(f_j(E) \ln |f_j(E)| - f_j(E)\right) + \varepsilon_j^{(0)}(E) \ln |f_j(E)| \mp \frac{\pi}{4} + O_E(k^{-1}).$$

The following proposition shows that the holonomy $\Theta(\zeta_E, k)$, which has a singular behavior as E tends to 0, can be regularized.

Proposition 6.9. *Fix a component* $p^{\pm} \in \mathfrak{C}^{\pm}$, *and let* $\zeta_E = [A_j(E), B_j(E)]$ *be a local path near the critical point* s_j. *Assume moreover that* ζ_E *is oriented according to the flow of* a_0. *Then there exists a sequence of* $\mathbb{R}/2\pi\mathbb{Z}$-*valued functions* $g_\zeta(\cdot, k) \in \mathscr{C}^{\infty}(D^{\pm})$, $E \mapsto g_{\zeta_E}(k)$, *admitting an asymptotic expansion in* $k \mathscr{C}^{\infty}(D^{\pm})[\![k^{-1}]\!]$ *of the form*

$$g_\zeta(E, k) = \sum_{\ell=-1}^{+\infty} k^{-\ell} g_\zeta^{(\ell)}(E),$$

such that

$$g_\zeta(E, k) = \Theta(\zeta_E, k) - i \ln v_j^{\pm}(E) \pmod{2\pi\mathbb{Z}} \quad \text{for all } E \in I^{\pm}.$$

The first terms of the asymptotic expansion of $g_\zeta(\cdot, k)$ *are given, for* $E \in I^{\pm}$, *by*

$$g_\zeta^{(-1)}(E) = c_0(\zeta_E) + \left(f_j(E) \ln |f_j(E)| - f_j(E)\right) + \Phi_{A_j(E)}^{(-1)}(A_j(E)) - \Phi_{B_j(E)}^{(-1)}(B_j(E)) \quad (17)$$

and

$$g_\zeta^{(0)}(E) = c_1(\zeta_E) + \mathrm{hol}_{\delta_{\zeta_E}}(\zeta_E) \mp \frac{\pi}{4} + \varepsilon_j^0(E) \ln |f_j(E)| + \Phi_{A_j(E)}^{(0)}(A_j(E)) - \Phi_{B_j(E)}^{(0)}(B_j(E)). \quad (18)$$

Proof. We can assume that the paths ζ_E, $E \in D^{\pm}$ all entirely lie in the open set Ω_{s_j} where the normal form of Proposition 5.1 is valid. Endow each edge e_i with the section $\phi_{k,E}^{(i)}$ defined earlier; by Lemma 6.7, these sections can be used to compute a new holonomy $\widetilde{\Theta}(\zeta_E, k)$. But we know how the different sections $\phi_{k,E}^{(i)}$ are related: (7) shows that $\widetilde{\Theta}(\zeta_E, k) - i \ln v_j^{\pm}(E) = 0$. Now, coming back to the microlocal solutions u_{A_j}, u_{B_j}, we have that $\Theta(\zeta_E, k) = \widetilde{\Theta}(\zeta_E, k) + c(E, k)$, where $c(E, k)$ admits an asymptotic expansion in $k \mathscr{C}^{\infty}(D^{\pm})[\![k^{-1}]\!]$. □

Since the sections $\phi_{k,E}^{(i)}$, $i = 1 \ldots 4$, form a standard basis at s_j, they can also be used to compute the holonomy $\mathrm{hol}_\mathfrak{L}$. Of course, for this choice of sections, one has $\mathrm{hol}_\mathfrak{L}(\zeta_0) = 1$. This allows to obtain the following result.

Proposition 6.10. *Let γ be a cycle in Γ_0, oriented according to the Hamiltonian flow of a_0, and of the form*

$$\gamma = \zeta_1^{\mathrm{loc}}(0)\zeta_1^{\mathrm{reg}}(0)\zeta_2^{\mathrm{loc}}(0)\zeta_2^{\mathrm{reg}}(0) \cdots \zeta_p^{\mathrm{loc}}(0)\zeta_p^{\mathrm{reg}}(0),$$

where ζ_j^{loc} and ζ_j^{reg} are local and regular paths, respectively, in the sense introduced earlier. Define

$$g(0, k) \sim \sum_{\ell=-1}^{+\infty} g^{(\ell)}(0)k^{-\ell}$$

as the sum

$$g(0, k) = g_{\zeta_1^{\mathrm{loc}}}(0, k) + g_{\zeta_1^{\mathrm{reg}}}(0, k) + \cdots + g_{\zeta_p^{\mathrm{loc}}}(0, k) + g_{\zeta_p^{\mathrm{reg}}}(0, k),$$

where $g_{\zeta_j^{\mathrm{loc}}}$ is given by Proposition 6.9 and $g_{\zeta_j^{\mathrm{reg}}}(E, k) = \Theta(\zeta_j^{\mathrm{reg}}(E), k)$. Then

$$\mathrm{hol}_\mathfrak{L}(\gamma) = \exp(ig(0, k)) + O(k^{-\infty}).$$

Proof. Notice that $\tilde{g}_{\zeta_j^{\mathrm{loc}}}(0, k) = 0$, where $\tilde{g}_{\zeta_j^{\mathrm{loc}}}(\,\cdot\,, k)$ is defined as $g_{\zeta_j^{\mathrm{loc}}}(\,\cdot\,, k)$, replacing $\Theta(\zeta_j^{\mathrm{loc}}, k)$ by $\widetilde{\Theta}(\zeta_j^{\mathrm{loc}}, k)$, and hence $\mathrm{hol}_\mathfrak{L}(\zeta_j^{\mathrm{loc}}(0)) = \exp(i\tilde{g}_{\zeta_j^{\mathrm{loc}}}(0, k))$. As in the previous proof, come back to the solutions u_{A_j}, u_{B_j}, and set

$$c_j(E, k) = g_{\zeta_j^{\mathrm{loc}}}(E, k) - \tilde{g}_{\zeta_j^{\mathrm{loc}}}(0, k).$$

Putting $\tilde{g}_{\zeta_j^{\mathrm{reg}}}(E, k) = \widetilde{\Theta}(\zeta_j^{\mathrm{reg}}(E), k)$, a simple computation shows that

$$\sum_{j=1}^{p} \tilde{g}_{\zeta_j^{\mathrm{reg}}}(E, k) = \sum_{j=1}^{p}(g_{\zeta_j^{\mathrm{reg}}}(E, k) + c_j(E, k)),$$

and the conclusion follows. $\square$

This is enough to prove the second part of Theorem 6.4, recalled in the following corollary.

Corollary 6.11. *The first two terms in the asymptotic expansion of the phase of $\mathrm{hol}_\mathfrak{L}(\gamma)$ are given by* (14).

Note that γ cannot always be obtained as a limit of smooth families of regular cycles; consider for instance the cycles $\gamma_1, \gamma_2, \gamma_3$ in the example treated in Section 7C (see Figures 13, 14). This is why the proof of this result requires some care.

Proof. We start with the case of a cycle γ oriented according to the Hamiltonian flow of a_0. Since $\mathfrak{e}_j^0(0) = 0$, formula (17) gives, for $j \in [\![1, p]\!]$,

$$g_{\zeta_j^{\mathrm{loc}}}^{(-1)}(0) = c_0(\zeta_j^{\mathrm{loc}}(0)) + \Phi_{A_j}^{(-1)}(A_j(0)) - \Phi_{B_j}^{(-1)}(B_j(0)),$$

while Proposition 6.6 shows that (identifying $j = p + 1$ with $j = 1$)

$$g_{\zeta_j^{\mathrm{reg}}}^{(-1)}(0) = c_0(\zeta_j^{\mathrm{reg}}(0)) + \Phi_{B_j}^{(-1)}(B_j(0)) - \Phi_{A_{j+1}}^{(-1)}(A_{j+1}(0)).$$

Consequently,

$$g^{(-1)}(0) = c_0(\gamma).$$

Let us now compute the subprincipal term $g^{(0)}_{\zeta^{\mathrm{loc}}_j}(0)$. Recall that it is equal to the limit of

$$c_1(\zeta^{\mathrm{loc}}_j(E)) + \mathrm{hol}_{\delta_{\zeta^{\mathrm{loc}}_j(E)}}(\zeta^{\mathrm{loc}}_j(E)) \mp \frac{\pi}{4} + \varepsilon^{(0)}_j(E)\ln|f_j(E)| + \Phi^{(0)}_{A_j(E)}(A_j(E)) - \Phi^{(0)}_{B_j(E)}(B_j(E))$$

as E goes to 0, which is equal to

$$\Phi^{(0)}_{A_j(0)}(A_j(0)) - \Phi^{(0)}_{B_j(0)}(B_j(0)) \mp \frac{\pi}{4} + \lim_{E\to 0}\big(c_1(\zeta^{\mathrm{loc}}_j(E)) + \mathrm{hol}_{\delta_{\zeta^{\mathrm{loc}}_j(E)}}(\zeta^{\mathrm{loc}}_j(E)) + \varepsilon^{(0)}_j(E)\ln|f_j(E)|\big).$$

First, we show that

$$\lim_{E\to 0}\big(c_1(\zeta^{\mathrm{loc}}_j(E)) + \varepsilon^{(0)}_j(E)\ln|f_j(E)|\big) = \tilde{c}_1(\zeta^{\mathrm{loc}}_j(0)), \tag{19}$$

where $\tilde{c}_1$ was introduced in Definition 6.2. Decompose

$$c_1(\zeta^{\mathrm{loc}}_j(E)) = \int_{\zeta^{\mathrm{loc}}_j(E)} \nu + \int_{\zeta^{\mathrm{loc}}_j(E)} \kappa_E,$$

where we recall that $-i\nu$ stands for the local connection 1-form associated to the Chern connection of L_1, and κ_E is such that $\kappa_E(X_{a_0}) = -a_1$. Of course, the term $\int_{\zeta^{\mathrm{loc}}_j(E)} \nu$ converges to $\int_{\zeta^{\mathrm{loc}}_j(0)} \nu$ as E tends to 0. Moreover, we have seen that there exist a symplectomorphism χ_E and a smooth function g^E_j such that $(g^E_j)'(0) > 0$ and

$$(a_0 \circ \chi_E^{-1})(x, \xi) - E = g^E_j(x\xi). \tag{20}$$

Hence, if we denote by $\tilde{a}_0$ (resp. $\tilde{a}_1$, $\tilde{\kappa}_E$) the pullback of a_0 (resp. a_1, κ_E) by χ_E^{-1}, we have

$$X_{\tilde{a}_0}(x, \xi) = (g^E_j)'(x\xi)X_{x\xi}(x, \xi),$$

so that $\tilde{\kappa}_E$ is characterized by

$$\tilde{\kappa}_E(X_{x\xi}) = \frac{-\tilde{a}_1(x, \xi)}{(g^E_j)'(x\xi)}.$$

Since $(g^E_j)'(0) \neq 0$, the function

$$b(x, \xi) = \frac{-\tilde{a}_1(x, \xi)}{(g^E_j)'(x\xi)}$$

is smooth (considering a smaller neighborhood of s_j for the definition of ζ^{loc}_j if necessary). Moreover, from (20), one finds that $(g^E_j)'(0) = |\det(\mathrm{Hess}(a_0)(s))|^{-1/2}$, which yields $b(0) = \varepsilon^{(0)}_j(0)$. Using a known result (see [Guillemin and Schaeffer 1977, Theorem 2, p. 175] for instance), we can construct smooth functions $F : \mathbb{R}^2 \to \mathbb{R}$ and $K : \mathbb{R} \to \mathbb{R}$ such that

$$b(x, \xi) = K(x\xi) - L_{X_{x\xi}}F(x, \xi);$$

since $x\xi = f_j(E)$ whenever $\chi_E^{-1}(x, \xi)$ belongs to Γ_E, this can be written

$$b(x, \xi) = K(f_j(E)) - L_{X_{x\xi}}F(x, \xi).$$

Therefore, the function

$$G = K(f_j(E)) \ln |x| - F \quad (\text{or } - K(f_j(E)) \ln |\xi| - F \text{ where } x = 0)$$

restricted to $\chi(\Gamma_E)$ is a primitive of $\tilde{\kappa}_E$. This yields

$$\int_{\zeta_j^{\text{loc}}(E)} \kappa_E = G(\tilde{B}_j) - G(\tilde{A}_j) = K(f_j(E))(\ln |x_{B_j}| - \ln |x_{A_j}|) + F((\tilde{A}_j) - F(\tilde{B}_j)), \qquad (21)$$

where $\tilde{m} = \chi_E(m)$ for any point $m \in M$, and (x_m, ξ_m) are the coordinates of $\tilde{m}$ (E being implicit to simplify notation). Writing $\ln |x_{B_j}| - \ln |x_{A_j}| = \ln |x_{B_j} \xi_{A_j}| - \ln |x_{A_j} \xi_{A_j}|$, we obtain

$$\int_{\zeta_j^{\text{loc}}(E)} \kappa_E + \varepsilon_j^{(0)}(E) \ln |f_j(E)|$$
$$= F(\tilde{A}_j) - F(\tilde{B}_j) + K(f_j(E)) \ln |x_{B_j} \xi_{A_j}| + (\varepsilon_j^{(0)}(E) - K(f_j(E))) \ln |f_j(E)|.$$

By definition of K, $b(0) - K(0) = 0$, hence $K(f_j(E)) = b(0) + O(f_j(E)) = \varepsilon_j^{(0)}(0) + O(f_j(E))$. Thus, the term $(\varepsilon_j^{(0)}(E) - K(f_j(E))) \ln |f_j(E)|$ tends to zero as E tends to zero; this induces

$$\lim_{E \to 0} \int_{\zeta_j^{\text{loc}}(E)} \kappa_E + \varepsilon_j^{(0)}(E) \ln |f_j(E)| = F(\tilde{A}_j) - F(\tilde{B}_j) + K(f_j(E)) \ln |x_{B_j} \xi_{A_j}|$$

(one must keep in mind that in this formula, we should write $\tilde{A}_j = \tilde{A}_j(0)$, etc.). Now, if a and b are points on $\zeta_j^{\text{loc}}(0)$ located in $[A_j, s_{m_j}]$ and $[s_{m_j}, B_j]$, respectively, then the term on the right-hand side of the previous equation is equal to

$$I = \lim_{a,b \to s_j} \left(F(\tilde{A}_j) - F(\tilde{a}) + F(\tilde{b}) - F(\tilde{B}_j) + K(f_j(E)) \ln |x_{B_j} \xi_{A_j}| \right).$$

Using (21), it is easily seen that

$$I = \lim_{a,b \to s_j} \left(\int_{[A_j, a]} \kappa_E + \int_{[b, B_j]} \kappa_E + \varepsilon_j^{(0)}(0) \ln |x_b \xi_a| \right).$$

Remembering Definition 6.2, this proves (19). Since $g_{\zeta_j^{\text{loc}}}^{(0)}$ and the quantities

$$\Phi_{A_j}^{(-1)}(A_j) - \Phi_{B_j}^{(-1)}(B_j) \quad \text{and} \quad \Phi_{A_j}^{(0)}(A_j) - \Phi_{B_j}^{(0)}(B_j)$$

are continuous at $E = 0$, the term

$$\text{hol}_{\delta_{\zeta_j^{\text{loc}}(E)}}(\zeta_j^{\text{loc}}(E))$$

is continuous at $E = 0$. Hence, if we sum up all the contributions from regular and local paths, we finally obtain

$$g^{(0)}(\gamma) = \tilde{c}_1(\gamma) + \sum_{m=1}^{p} \frac{\rho_m \pi}{4} + \ell(\gamma),$$

where ρ_m and $\tilde{c}_1$ were introduced in Definitions 6.3 and 6.2, respectively, and $\ell(\gamma)$ is the quantity

$$\ell(\gamma) = \sum_{j=1}^{p} \left(\text{hol}_{\delta_{\zeta_j^{\text{reg}}(0)}}(\zeta_j^{\text{reg}}(0)) + \lim_{E \to 0} \text{hol}_{\delta_{\zeta_j^{\text{loc}}(E)}}(\zeta_j^{\text{loc}}(E)) \right);$$
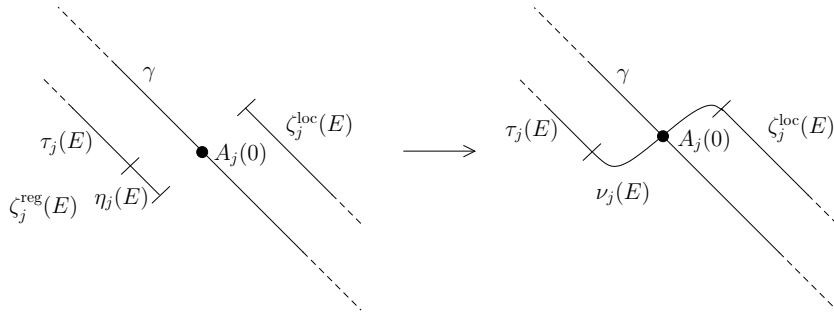
Figure 3. Computation of $\ell(\gamma)$.

it is not hard to show that $\ell(\gamma)$ is independent of the choice of the local and regular paths. Furthermore, let ϵ be the index of any smooth embedded cycle which is a continuous deformation of γ. If the regular and local paths can be chosen so that they all lie in the same connected component γ_E of Γ_E, it is clear that $\ell(\gamma) = \epsilon$, because for $E \neq 0$,

$$\sum_{j=1}^{p} \left(\mathrm{hol}_{\delta_{\zeta_j^{\mathrm{reg}}(E)}}(\zeta_j^{\mathrm{reg}}(E)) + \mathrm{hol}_{\delta_{\zeta_j^{\mathrm{loc}}(E)}}(\zeta_j^{\mathrm{loc}}(E)) \right) = \epsilon(\gamma_E) = \epsilon.$$

If it is not the case, we remove a small path $\eta_j(E)$ of $\zeta_j^{\mathrm{reg}}(E)$ at any point A_j or B_j where there is a change of connected component, and replace it by a smooth path $\nu_j(E)$ connecting $\zeta_j^{\mathrm{reg}}(E)$ and $\zeta_j^{\mathrm{loc}}(E)$ (see Figure 3). We obtain a smooth path $\tilde{\gamma}(E)$; on the one hand, one has $\epsilon(\tilde{\gamma}(E)) = \epsilon$. On the other hand, $\epsilon(\tilde{\gamma}(E))$ is the sum of the holonomies of the paths composing $\tilde{\gamma}(E)$. But, if we denote by $\tau_j(E)$ the part of $\zeta_j^{\mathrm{reg}}(E)$ that remains when we remove $\eta_j(E)$, we have

$$\mathrm{hol}_{\delta_{\tau_j(E)}}(\tau_j(E)) = \mathrm{hol}_{\delta_{\zeta_j^{\mathrm{reg}}(E)}}(\zeta_j^{\mathrm{reg}}(E)) - \mathrm{hol}_{\delta_{\eta_j(E)}}(\eta_j(E)),$$

which implies

$$\mathrm{hol}_{\delta_{\tau_j(E)}}(\tau_j(E)) + \mathrm{hol}_{\delta_{\nu_j(E)}}(\nu_j(E)) \xrightarrow[E \to 0]{} \mathrm{hol}_{\delta_{\zeta_j^{\mathrm{reg}}(0)}}(\zeta_j^{\mathrm{reg}}(0))$$

because

$$\mathrm{hol}_{\delta_{\nu_j(E)}}(\nu_j(E)) - \mathrm{hol}_{\delta_{\eta_j(E)}}(\eta_j(E)) \xrightarrow[E \to 0]{} 0.$$

This shows that $\ell(\gamma) = \epsilon$, which concludes the proof for this first case, where γ is oriented according to the Hamiltonian flow of a_0.

If the orientation of the cycle γ is opposite to the one of the flow of X_{a_0}, we only have to change the sign of the holonomy.

It remains to investigate the case where there are some paths in γ oriented according to the flow of X_{a_0} and some oriented in the opposite direction, which means γ is smooth at some critical point s. We can use the analysis above by introducing two local paths ζ_1^{loc} and ζ_2^{loc} at s as in Figure 4 (we make a small move forwards and backwards on an edge added to γ); one can obtain the claimed result by looking carefully at the obtained holonomies, remembering that the two paths have opposite orientation on the added edge. Note that the choice of the added edge does not change the result. $\qquad \square$
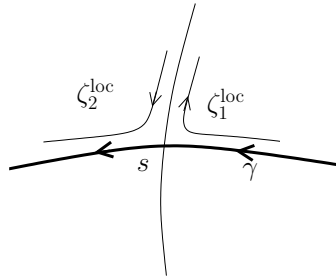
Figure 4. Case of a cycle γ, smooth at s.

6D. *Derivation of the Bohr–Sommerfeld conditions.* The previous results allow to compute the spectrum of A_k in an interval of size $O(1)$ around the singular energy. Indeed, let γ_E, $E \in I^{\pm}$ be a connected component of the level $a_0^{-1}(E)$ and γ be the cycle in Γ_0 obtained by letting E go to 0. Then one can choose the local and regular paths used to compute the holonomy $\mathrm{hol}_{\mathfrak{L}}(\gamma)$ so that they all lie on γ_E, and define $g(E, k)$ as the sum

$$g(E, k) = g_{\zeta_1^{\mathrm{loc}}}(E) + g_{\zeta_1^{\mathrm{reg}}}(E) + \cdots + g_{\zeta_p^{\mathrm{loc}}}(E) + g_{\zeta_p^{\mathrm{reg}}}(E).$$

Furthermore, the matrix of change of basis associated to the sections $\phi_{k,E}^{(i)}$ is given by

$$T_j(E) = \exp\left(-\frac{i\pi}{4}\right) \mathscr{E}_k(k\mathfrak{e}_j(E, k)) \begin{pmatrix} 1 & i \exp(-k\pi\mathfrak{e}_j(E, k)) \\ i \exp(-k\pi\mathfrak{e}_j(E, k)) & 1 \end{pmatrix},$$

where the function $\mathscr{E}_k$ is defined in (9). To compute eigenvalues near E, apply Theorem 6.1 with T_j replaced by $T_j(E)$ and $\mathrm{hol}_{\mathfrak{L}}(\gamma)$ by $\exp(ig(E, k))$. Applying Stirling's formula, we obtain

$$T_j(E) = \exp(ik\theta(E, k)) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + O(k^{-1}), \quad f_j(E) > 0,$$

and

$$T_j(E) = \exp(ik\theta(E, k)) \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + O(k^{-1}), \quad f_j(E) < 0,$$

with $\theta(E, k) = f_j(E) \ln|f_j(E)| - f_j(E) + k^{-1}(\varepsilon_j^{(0)}(E) \ln|f_j(E)| - \pi/4)$. Together with equations (17) and (18), this ensures that we recover the usual Bohr–Sommerfeld conditions away from the critical energy.

In the rest of the paper, we will look for eigenvalues of the form $k^{-1}e + O(k^{-2})$, where e is allowed to vary in a compact set. Hence, we have to replace A_k by $A_k - k^{-1}e$; this operator still has principal symbol a_0, but its subprincipal symbol is $a_1 - e$. Thanks to Theorem 6.4, we are able to compute the singular holonomy and the invariants ε_j up to $O(k^{-2})$; hence, we approximate the spectrum up to an error of order $O(k^{-2})$.

6E. *The case of a unique saddle point.* If Γ_0 contains a unique saddle point, it is not difficult to write the Bohr–Sommerfeld conditions in a more explicit form. The critical level Γ_0 looks like a figure eight. We choose the convention for the cut edges and cycles as in Figure 5.
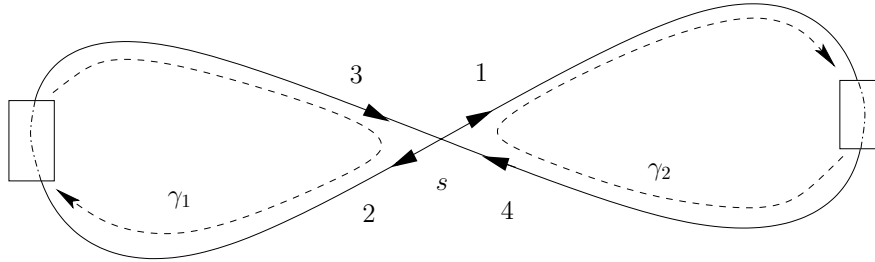
Figure 5. The singular level $\Gamma_0 = a_0^{-1}(0)$ and the choice of cut edges and cycles.

Let s be the saddle point, and let $\varepsilon(e, k)$ be the invariant associated to the operator $A_k - k^{-1}e$ at s; one has $\varepsilon^{(0)}(e) = \varepsilon^{(0)}(0) + e\,|\det(\mathrm{Hess}(a_0)(s))|^{-1/2}$. Denote by $h_j(e, k) = \exp(i\theta_j(e, k))$ the holonomy of the loop γ_j in $\mathfrak{L}$; remember that θ_j is given by

$$\theta_j(e, k) = k\, c_0(\gamma_j) + \tilde{c}_1(\gamma_j) + \tilde{\epsilon}(\gamma_j)\pi + O(k^{-1}).$$

The Bohr–Sommerfeld conditions are given by the holonomy equations

$$x_4 = h_2 x_1, \quad x_3 = h_1 x_2,$$

and by the transfer relation at the critical point

$$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where $T = T(\varepsilon)$ is defined in (8). Using Lemma 2 of [Colin de Verdière and Parisse 1994b], the quantization rule can in fact be written as a real scalar equation.

Proposition 6.12. *The equation $A_k u_k = k^{-1}e\, u_k + O(k^{-\infty})$ has a normalized eigenfunction if and only if e satisfies the condition*

$$\frac{1}{\sqrt{1 + \exp(2\pi\varepsilon)}}\cos\left(\frac{\theta_1 - \theta_2}{2}\right) = \sin\left(\frac{\theta_1 + \theta_2}{2} + \frac{\pi}{4} + \varepsilon \ln k - \arg\Gamma\left(\tfrac{1}{2} + i\varepsilon\right)\right), \qquad (22)$$

where we wrote for the sake of brevity θ_j, ε instead of $\theta_j(e, k)$, $\varepsilon(e, k)$ (see definitions above).

7. Examples

We conclude by investigating two examples on the torus and one on the sphere; these examples present various topologies. More precisely, using the terminology of [Oshemkov 1994; Bolsinov and Fomenko 2004] for atoms (neighborhoods of singular levels of Morse functions), we provide an example of a type B atom — the only type in complexity 1 (here, complexity means the number of critical points on the singular level) in the orientable case — and two examples of atoms of complexity 2: one is of type C_2 (xy on the sphere S^2) and the other is of type C_1 (Harper's Hamiltonian on the torus $\mathbb{T}^2$). It is a remarkable fact that these two examples are natural not only as the canonical realization of the atom on a surface but also because they come from the simplest possible Toeplitz operators with critical level of given type.

Note that there are two other types of atoms of complexity 2 in the orientable case (more precisely, types D_1 and D_2); it would be interesting to realize each of them as a hyperbolic level of the principal symbol of a selfadjoint Toeplitz operator and to complete this study. Note that in the context of pseudodifferential operators, Colin de Verdière and Parisse [1999] treated the case of a type D_1 atom (the triple well potential) among some other examples. More generally, one could use the classification of Bolsinov, Fomenko and Oshemkov to write the Bohr–Sommerfeld conditions for all cases in low complexity (≤ 3 for instance); however, the case of two critical points already gives rise to rather tedious computations.

The details of the quantization of the torus and the sphere are quite standard. Nevertheless, for the sake of completeness, we will recall a few of them at the beginning of each paragraph.

7A. *Height function on the torus.* Firstly, we consider the quantization of the height function on the torus. This is one of the first examples in Morse theory, perhaps because this is the simplest and most intuitive example with critical points of each type. In particular, the description of the two hyperbolic levels is quite simple.

Endow $\mathbb{R}^2$ with the linear symplectic form ω_0 and let $L_0 \to \mathbb{R}^2$ be the complex line bundle with Hermitian form and connection defined in Section 2D3. Let K be the canonical line of $\mathbb{R}^2$ with respect to its standard complex structure j: $K = \{\alpha \in (\mathbb{R}^2)^* \otimes \mathbb{C} \mid \alpha(j \cdot) = i\alpha\}$. Choose a half-form line, that is, a complex line δ with an isomorphism $\varphi : \delta^{\otimes 2} \to K$. There is a natural scalar product on K such that the square of the norm of α is $i\alpha \wedge \bar{\alpha}/\omega_0$; endow δ with the scalar product $\langle \cdot, \cdot \rangle_\delta$ such that φ is an isometry. The half-form bundle we work with, that we still denote by δ, is the trivial line bundle with fiber δ over $\mathbb{R}^2$.

Consider a lattice Λ with symplectic volume 4π. The Heisenberg group $H = \mathbb{R}^2 \times U(1)$ with product

$$(x, u) \cdot (y, v) = \left(x + y, uv \exp\left(\frac{i}{2}\omega_0(x, y)\right)\right)$$

acts on the bundle $L_0 \to \mathbb{R}^2$, with action given by the same formula. This action preserves the prequantum data, and the lattice Λ injects into H; therefore, the fiber bundle L_0 reduces to a prequantum bundle L over $\mathbb{T}^2 = \mathbb{R}^2/\Lambda$. The action extends to the fiber bundle L_0^k by

$$(x, u) \cdot (y, v) = \left(x + y, u^k v \exp\left(\frac{ik}{2}\omega_0(x, y)\right)\right).$$

We let the Heisenberg group act trivially on δ. We obtain a half-form bundle $\tilde{\delta}$ over $\mathbb{T}^2$ and an action

$$T^* : \Lambda \to \mathrm{End}(\mathscr{C}^\infty(\mathbb{R}^2, L_0^k \otimes \delta)), \quad u \mapsto T_u^*.$$

The Hilbert space $\mathscr{H}_k = H^0(M, L^k \otimes \tilde{\delta})$ can naturally be identified with the space $\mathscr{H}_{\Lambda,k}$ of holomorphic sections of $L_0^k \otimes \delta \to \mathbb{R}^2$ which are invariant under the action of Λ, endowed with the Hermitian product

$$\langle \varphi, \psi \rangle = \int_D \langle \varphi, \psi \rangle_\delta \, |\omega_0|,$$

where D is the fundamental domain of the lattice. Furthermore, $\Lambda/2k$ acts on $\mathscr{H}_{\Lambda,k}$. Let e and f be generators of Λ satisfying $\omega_0(e, f) = 4\pi$; one can show that there exists an orthonormal basis $(\psi_\ell)_{\ell \in \mathbb{Z}/2k\mathbb{Z}}$
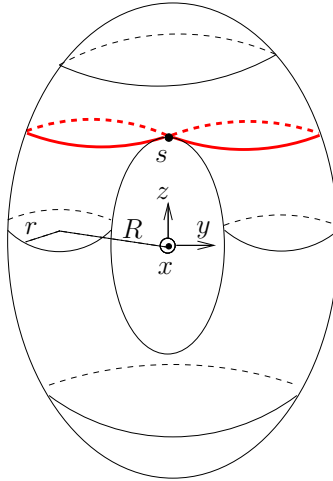
Figure 6. Height function on the torus.

of $\mathcal{H}_{\Lambda,k}$ such that

$$\left. \begin{array}{l} T^*_{e/2k}\psi_\ell = w^\ell \psi_\ell \\ T^*_{f/2k}\psi_\ell = \psi_{\ell+1} \end{array} \right\} \quad \text{for all } \ell \in \mathbb{Z}/2k\mathbb{Z},$$

with $w = \exp(i\pi/k)$. The sections ψ_ℓ can be expressed in terms of Θ functions.

Set $M_k = T^*_{e/2k}$ and $L_k = T^*_{f/2k}$. Let (q, p) be coordinates on $\mathbb{R}^2$ associated to the basis (e, f) and $[q, p]$ be the equivalence class of (q, p). Both M_k and L_k are Toeplitz operators, with respective principal symbols $[q, p] \mapsto \exp(2i\pi p)$ and $[q, p] \mapsto \exp(2i\pi q)$, and vanishing subprincipal symbols. For more details, see for instance [Charles and Marché 2011, §2.2, §3.1].

It is a well-known fact that $\mathbb{T}^2$ is diffeomorphic to the surface shown in Figure 6, which is obtained by rotating a circle of radius r around a circle of radius $R > r$ contained in the yz plane; the diffeomorphism is given by the explicit formulas

$$x = r\sin(2\pi q), \quad y = (R + r\cos(2\pi q))\cos(2\pi p), \quad z = (R + r\cos(2\pi q))\sin(2\pi p).$$

Hence, the Hamiltonian that we consider is

$$a_0(q, p) = (R + r\cos(2\pi q))\sin(2\pi p)$$

on the fundamental domain D. We try to quantize it, that is, find a selfadjoint Toeplitz operator A_k with principal symbol a_0. The Toeplitz operators

$$B_k = \frac{1}{2i}(M_k - M_k^*), \quad C_k = R\Pi_k + \frac{r}{2}(L_k + L_k^*),$$

are selfadjoint, and

$$\sigma_{\text{norm}}(B_k) = \sin(2\pi p) + O(\hbar^2), \quad \sigma_{\text{norm}}(C_k) = R + r\cos(2\pi q) + O(\hbar^2).$$

Hence $A_k = \frac{1}{2}(B_k C_k + C_k B_k)$ is a selfadjoint Toeplitz operator with normalized symbol $a_0 + O(\hbar^2)$. Its matrix in the basis $(\psi_\ell)_{\ell \in \mathbb{Z}/2k\mathbb{Z}}$ is written as

$$
\begin{pmatrix}
R\alpha_0 & \frac{r}{4}(\alpha_0 + \alpha_1) & 0 & \cdots & 0 & \frac{r}{4}(\alpha_{2k-1} + \alpha_0) \\
\frac{r}{4}(\alpha_0 + \alpha_1) & R\alpha_1 & \frac{r}{4}(\alpha_1 + \alpha_2) & 0 & \cdots & 0 \\
0 & \frac{r}{4}(\alpha_1 + \alpha_2) & R\alpha_2 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & R\alpha_{2k-2} & \frac{r}{4}(\alpha_{2k-2} + \alpha_{2k-1}) \\
\frac{r}{4}(\alpha_0 + \alpha_{2k-1}) & 0 & \cdots & 0 & \frac{r}{4}(\alpha_{2k-2} + \alpha_{2k-1}) & R\alpha_{2k-1}
\end{pmatrix}
\tag{23}
$$

with $\alpha_\ell = \sin(\ell\pi/k)$.

The level $\Gamma_{R-r} = a_0^{-1}(R-r)$ contains one hyperbolic point $s = \left(\frac{1}{2}, \frac{1}{4}\right)$. It is the union of the two branches

$$
p = \frac{1}{2\pi} \arcsin \frac{R-r}{R+r\cos(2\pi q)} \quad \text{and} \quad p = \frac{1}{2} - \frac{1}{2\pi} \arcsin \frac{R-r}{R+r\cos(2\pi q)}.
$$

The Hamiltonian vector field associated to a_0 is given by

$$
X_{a_0}(q, p) = \frac{1}{2}(R + r\cos(2\pi q)) \cos(2\pi p) \frac{\partial}{\partial q} + \frac{r}{2} \sin(2\pi q) \sin(2\pi p) \frac{\partial}{\partial p}.
$$

Moreover, one has

$$
\varepsilon^{(0)} = \frac{e}{\pi \sqrt{r(R-r)}}.
\tag{24}
$$

We choose the cycles γ_1 and γ_2 with the convention given in Section 6E. We have to compute the principal and subprincipal actions of γ_1, γ_2 and their indices $\tilde{\epsilon}$. Let us detail the calculations in the case of γ_1.

We parametrize γ_1 by

$$
q \mapsto \left(q, \frac{1}{2} - \frac{1}{2\pi} \arcsin \frac{R-r}{R+r\cos(2\pi q)}\right).
$$

The principal action is given by

$$
c_0(\gamma_1) = 2I(R, r) - 2\pi,
\tag{25}
$$

where $I(R, r)$ is the integral

$$
I(R, r) = \int_0^1 \arcsin \frac{R-r}{R+r\cos(2\pi q)} \, dq;
$$

unfortunately, we do not know any explicit expression for this integral, so for numerical computations, once the radii R and r are fixed, we obtain the value of $I(R, r)$ thanks to numerical integration routines.

On γ_1, the subprincipal form reads

$$
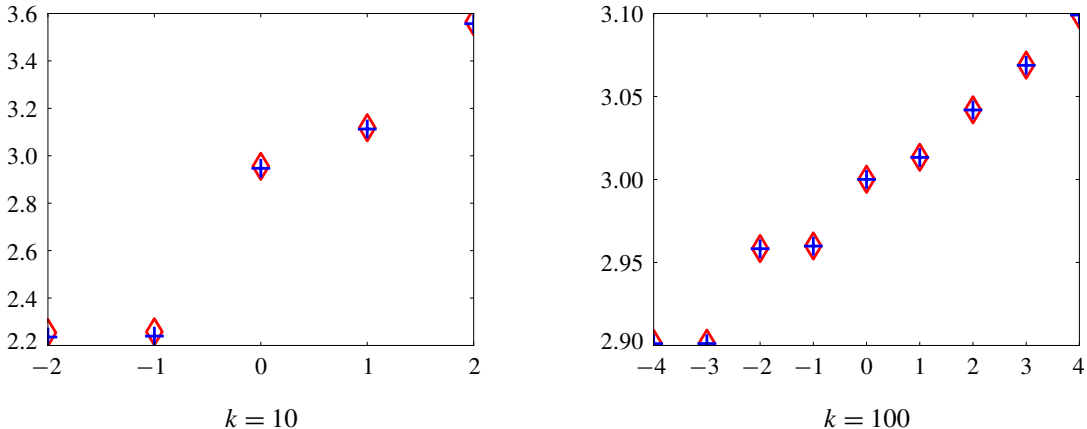\kappa_0 = \frac{-2e \, dq}{\sqrt{(R + r\cos(2\pi q))^2 - (R-r)^2}}.
$$

Figure 7. Eigenvalues in $[R-r-10k^{-1}, R-r+10k^{-1}]$; in red diamonds, the eigenvalues of A_k obtained numerically; in blue crosses, the theoretical eigenvalues derived from the Bohr–Sommerfeld conditions. The results are indexed with respect to the eigenvalue closest to the critical energy, labeled as 0. Even for $k = 10$, the method is very accurate.

One can obtain an explicit primitive thanks to any computer algebra system. Furthermore, some computations show that the symplectic area of the parallelogram $R_{a,b}$ is equal to

$$\int_{R_{a,b}} \omega = 8\pi \sqrt{\frac{r}{R-r}}\left(q_a - \tfrac{1}{2}\right)\left(\tfrac{1}{2} - q_b\right).$$

This yields the following value for the subprincipal action:

$$\tilde{c}_1(\gamma_1) = \varepsilon^{(0)} \ln\left(\frac{32}{\pi}\sqrt{\frac{r}{R}\left(1 - \frac{r}{R}\right)}\right). \tag{26}$$

Finally, the index associated to half-forms is $\tilde{\epsilon}(\gamma_1) = \tfrac{1}{4}$. For γ_2, one can check that

$$c_0(\gamma_2) = 2I(R, r), \quad \tilde{c}_1(\gamma_2) = \varepsilon^{(0)} \ln\left(\frac{32}{\pi}\sqrt{\frac{r}{R}\left(1 - \frac{r}{R}\right)}\right), \quad \tilde{\epsilon}(\gamma_2) = \tfrac{1}{4}. \tag{27}$$

With this data, one can test the Bohr–Sommerfeld condition for different pairs (R, r). We illustrate this with $(R, r) = (4, 1)$ (note that we have tested several pairs). We compare the eigenvalues obtained numerically from the matrix (23) and the ones derived from the Bohr–Sommerfeld conditions (22) in the interval $I = [R - r - 10k^{-1}, R - r + 10k^{-1}]$. In Figure 7, we plot the theoretical and numerical eigenvalues; Figure 8 shows the error between the eigenvalues and the solutions of the Bohr–Sommerfeld conditions for fixed k, while Figure 9 is a graph of the logarithm of the maximal error in the interval I as a function of $\ln k$.

7B. xy on the 2-sphere. Let us consider another simple example, but this time with two saddle points on the critical level. We will quantize the Hamiltonian $a_0(x, y, z) = xy$ on the sphere S^2. Let us briefly recall the details of the quantization of this surface.
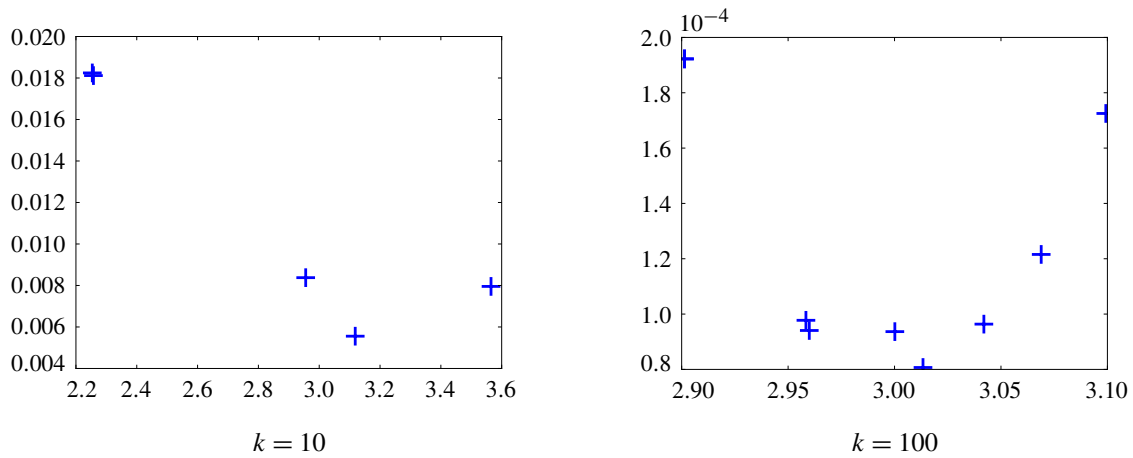
Figure 8. Absolute value of the difference between the numerical and theoretical eigenvalues; the error is smaller near the critical energy ($R - r = 3$ in this case).
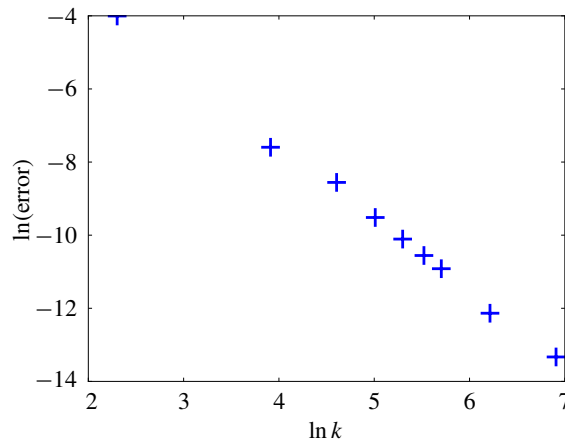
Figure 9. Logarithm of the maximal error as a function of the logarithm of k; the error displays a behavior in $O(k^{-2})$, as expected.

Start from the complex projective plane $\mathbb{CP}^1$ and let $L = \mathbb{O}(1)$ be the dual bundle of the tautological bundle

$$\mathbb{O}(-1) = \{(u, v) \in \mathbb{CP}^1 \times \mathbb{C}^2 \mid v \in u\},$$

with natural projection. L is a Hermitian, holomorphic line bundle; let us denote by ∇ its Chern connection. The 2-form $\omega = i \operatorname{curv}(\nabla)$ is the symplectic form on $\mathbb{CP}^1$ associated with the Fubini–Study Kähler structure, and $L \to \mathbb{CP}^1$ is a prequantum bundle. Moreover, the canonical bundle is naturally identified with $\mathbb{O}(-2)$, hence one can choose the line bundle $\delta = \mathbb{O}(-1)$ as a half-form bundle. The state space $\mathcal{H}_k = H^0(\mathbb{CP}^1, L^k \otimes \delta)$ can be identified with the space $\mathbb{C}_{p_k}[z_1, z_2]$ of homogeneous polynomials

of degree $p_k = k - 1$ in two variables. The polynomials

$$P_\ell(z_1, z_2) = \sqrt{\frac{(p_k + 1)\binom{p_k}{\ell}}{2\pi}} \, z_1^\ell z_2^{p_k - \ell}, \quad 0 \le \ell \le p_k,$$

form an orthonormal basis of $\mathcal{H}_k$. The sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ is diffeomorphic to $\mathbb{CP}^1$ via the stereographic projection (from the north pole to the plane $z = 0$). The symplectic form ω on $\mathbb{CP}^1$ is carried to the symplectic form $\omega_{S^2} = -\frac{1}{2}\Omega$, with Ω the usual area form on S^2 (the one which gives the area 4π). The operator A_k acting on the basis $(P_\ell)_{0 \le \ell \le p_k}$ by

$$A_k P_\ell = \frac{i}{p_k^2}(\alpha_{\ell,k} P_{\ell-2} - \beta_{\ell,k} P_{\ell+2}),$$

with

$$\alpha_{\ell,k} = \sqrt{\ell(\ell - 1)(p_k - \ell + 1)(p_k - \ell + 2)}$$

and

$$\beta_{\ell,k} = \sqrt{(\ell + 1)(\ell + 2)(p_k - \ell - 1)(p_k - \ell)},$$

is a Toeplitz operator with principal symbol $a_0(x, y, z) = xy$ and vanishing subprincipal symbol (for more details, one can consult [Bloch et al. 2003, §3] for instance). Note that $\alpha_{\ell,k} = \beta_{p_k-\ell,k}$, which implies that if λ is an eigenvalue of A_k, then $-\lambda$ is also.

The level $a_0^{-1}(0)$ is critical, and contains two saddle points: the poles N (north) and S (south). It is the union of the two great circles $x = 0$ and $y = 0$. We choose the cut edges and cycles as indicated in Figure 10.
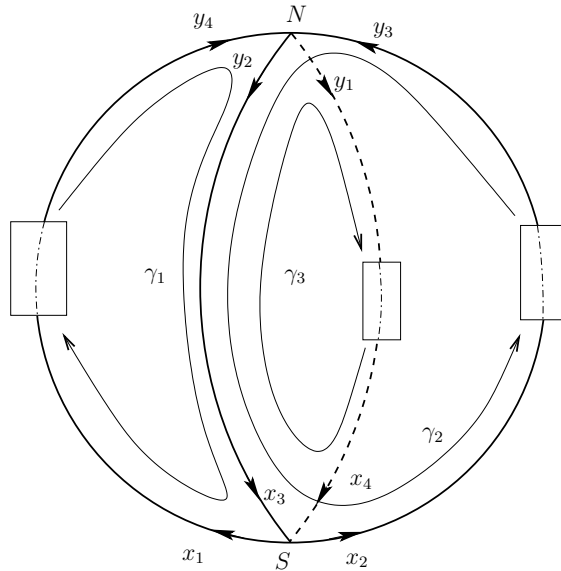

Figure 10. Choice of the cycles and cut edges.

Set $h_j = \mathrm{hol}_{\mathfrak{L}}(\gamma_j) = \exp(i\theta_j)$; remember that $\theta_j = kc_0(\gamma_j) + \tilde{c}_1(\gamma_j) + \tilde{\epsilon}(\gamma_j)\pi + O(k^{-1})$. The holonomy equations read

$$y_2 = x_3, \quad y_4 = h_1 x_1, \quad y_3 = h_2 x_2, \quad x_4 = h_3 y_1, \tag{28}$$

while the transfer equations are given by

$$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = T_S \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} y_3 \\ y_4 \end{pmatrix} = T_N \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{29}$$

The system (28) + (29) has a solution if and only if the matrix

$$U = T_S \begin{pmatrix} 0 & \exp(-i\theta_1) \\ \exp(-i\theta_2) & 0 \end{pmatrix} T_N \begin{pmatrix} 0 & \exp(-i\theta_3) \\ 1 & 0 \end{pmatrix}$$

admits 1 as an eigenvalue. The matrix U is unitary, and if we write $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, a straightforward computation shows that

$$|a|^2 = |d|^2 = \frac{1 - 2\cos(\theta_2 - \theta_1)\exp(-\pi(\varepsilon_S + \varepsilon_N)) + \exp(-2\pi(\varepsilon_S + \varepsilon_N))}{(1 + \exp(-2\pi\varepsilon_S))(1 + \exp(-2\pi\varepsilon_N))};$$

hence, by Lemma 2 of [Colin de Verdière and Parisse 1994b], 1 is an eigenvalue of U if and only if

$$|a| \sin\left(\frac{\arg(ad) - \pi}{2} - \arg a\right) = \sin\left(\frac{\arg(ad) - \pi}{2}\right).$$

This amounts to the equation

$$|a| \cos\left(\frac{\arg z - \arg w}{2}\right) = \sin\left(\frac{\arg z + \arg w}{2} + \arg\Gamma\left(\tfrac{1}{2} + i\varepsilon_N\right) + \arg\Gamma\left(\tfrac{1}{2} + i\varepsilon_S\right) - (\varepsilon_S + \varepsilon_N)\ln k\right),$$

with

$$z = \exp(-i(\theta_2 + \theta_3)) - \exp(-\pi(\varepsilon_S + \varepsilon_N) - i(\theta_1 + \theta_3))$$

and

$$w = \exp(-i\theta_1) - \exp(-\pi(\varepsilon_N + \varepsilon_S) - i\theta_2).$$

One has

$$\varepsilon_S^{(0)} = \varepsilon_N^{(0)} = \varepsilon^{(0)} = \frac{e}{2}. \tag{30}$$

Moreover, the principal actions are

$$c_0(\gamma_1) = -\frac{\pi}{2}, \quad c_0(\gamma_2) = \frac{\pi}{2}, \quad c_0(\gamma_3) = \pi. \tag{31}$$

Then, one finds for the subprincipal actions

$$\tilde{c}_1(\gamma_1) = 2\varepsilon^{(0)}\ln 2 = \tilde{c}_1(\gamma_2), \quad \tilde{c}_1(\gamma_3) = 0. \tag{32}$$

Finally, the indices $\tilde{\epsilon}$ are

$$\tilde{\epsilon}(\gamma_1) = \tfrac{3}{2}, \quad \tilde{\epsilon}(\gamma_2) = \tfrac{1}{2}, \quad \tilde{\epsilon}(\gamma_3) = 1. \tag{33}$$
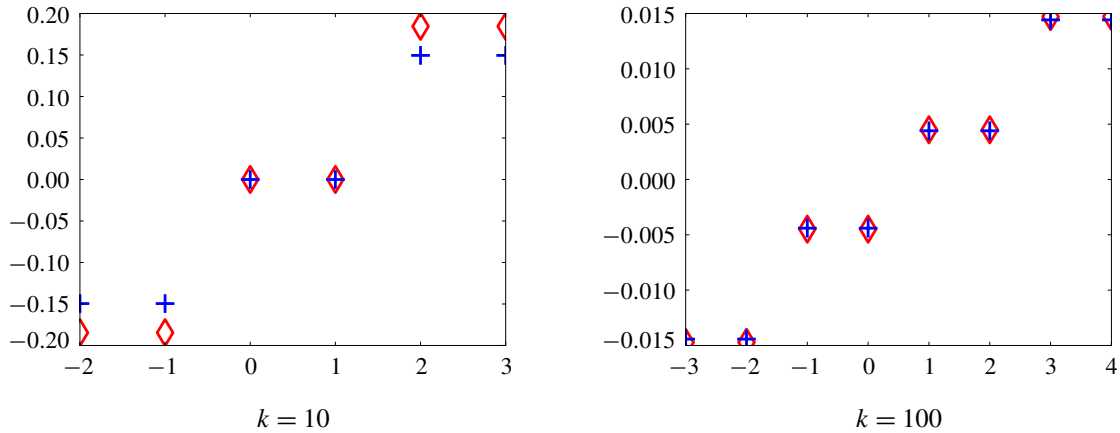
Figure 11. Eigenvalues in $[-2k^{-1}, 2k^{-1}]$; in red diamonds, the eigenvalues of A_k obtained numerically; in blue crosses, the theoretical eigenvalues derived from the Bohr–Sommerfeld conditions.

Figure 11 shows the theoretical eigenvalues obtained by using these results, as well as the numerical evaluation of the eigenvalues of A_k from its matrix form.

7C. *Harper's Hamiltonian on the torus.* Keeping the conventions and notation of the first example, we consider the Hamiltonian (sometimes known as Harper's Hamiltonian since it is related to Harper's equation [Helffer and Sjöstrand 1988])

$$a_0(q, p) = 2(\cos(2\pi p) + \cos(2\pi q))$$

on the torus. The operator $A_k = M_k + M_k^* + L_k + L_k^*$ is a Toeplitz operator with principal symbol a_0 and vanishing subprincipal symbol. Its matrix in the basis $(\psi_\ell)_{\ell \in \mathbb{Z}/2k\mathbb{Z}}$ is

$$\begin{pmatrix} 2\alpha_0 & 1 & 0 & \ldots & 0 & 1 \\ 1 & \ddots & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & \ddots & 1 \\ 1 & 0 & \ldots & 0 & 1 & 2\alpha_{2k-1} \end{pmatrix}$$

where

$$\alpha_\ell = \cos\frac{\ell\pi}{k}, \quad 0 \leq \ell \leq 2k - 1.$$

The critical level $\Gamma_0 = a_0^{-1}(0)$ contains two hyperbolic points: $s_1 = (0, \frac{1}{2})$ and $s_2 = (\frac{1}{2}, 0)$. On the fundamental domain, it is the union of the four segments described in Figure 12; hence, its image on the torus it is the union of two circles that intersect at two points.
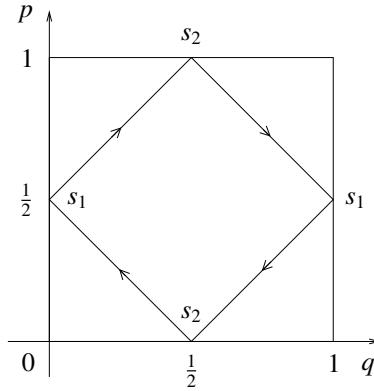
Figure 12. Critical level Γ_0 on the fundamental domain; the arrows indicate the direction of the Hamiltonian flow of a_0.
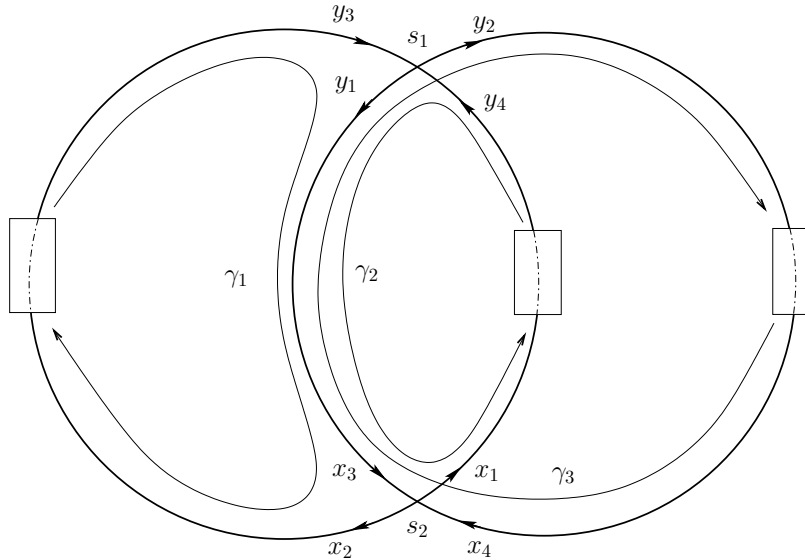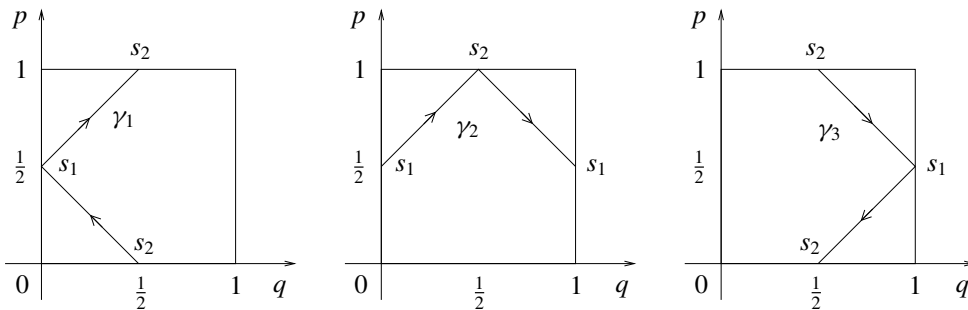

Figure 13. Choice of the cycles and cut edges.

Figure 14. Cycles on the fundamental domain.
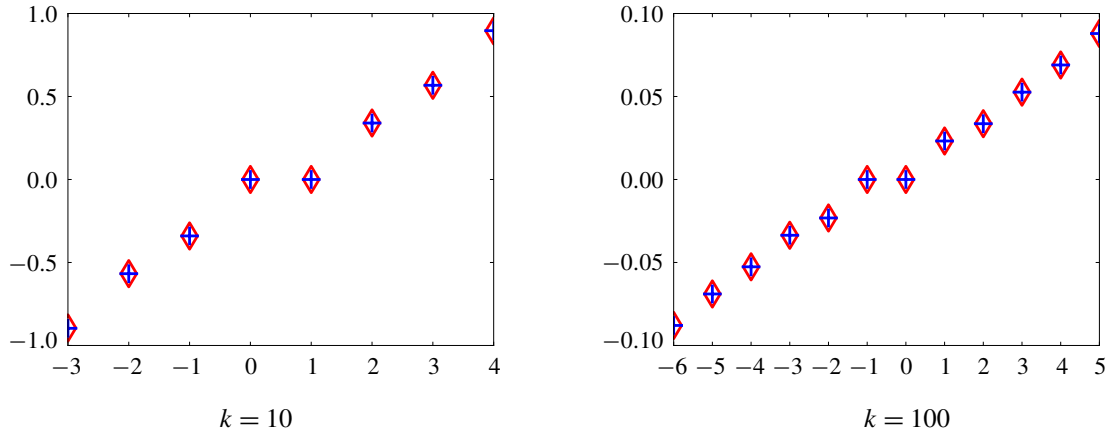
Figure 15. Eigenvalues in $[-10k^{-1}, 10k^{-1}]$; in red diamonds, the eigenvalues of A_k obtained numerically; in blue crosses, the theoretical eigenvalues derived from the Bohr–Sommerfeld conditions.

We choose the cycles and cut edges as in Figure 13 (for a representation of the two circles in a two-dimensional view) and Figure 14 (for a representation of the cycles on the fundamental domain).

We write the holonomy equations

$$y_1 = x_3, \quad y_3 = h_1 x_2, \quad y_4 = h_2 x_1, \quad x_4 = h_3 y_1, \tag{34}$$

and the transfer equations

$$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = T_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} y_3 \\ y_4 \end{pmatrix} = T_1 \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{35}$$

where $h_j = \mathrm{hol}_{\mathfrak{L}}(\gamma_j) = \exp(i\theta_j)$. Following the same steps as in the previous example, one can show that the system (34) + (35) has a solution if and only if e is a solution of the scalar equation

$$|a| \cos\left(\frac{\arg w - \arg z}{2}\right) = \cos\left(\frac{\arg z + \arg w}{2} + \arg \Gamma\left(\tfrac{1}{2} + i\varepsilon_1\right) + \arg \Gamma\left(\tfrac{1}{2} + i\varepsilon_2\right) - (\varepsilon_1 + \varepsilon_2) \ln k\right),$$

with

$$|a|^2 = \frac{\exp(-2\pi\varepsilon_1) + \exp(-2\pi\varepsilon_2) + 2\cos(\theta_2 - \theta_1)\exp(-\pi(\varepsilon_1 + \varepsilon_2))}{(1 + \exp(-2\pi\varepsilon_1))(1 + \exp(-2\pi\varepsilon_2))},$$

$$w = \exp(-\pi\varepsilon_2 - i(\theta_2 + \theta_3)) + \exp(-\pi\varepsilon_1 - i(\theta_1 + \theta_3)),$$

and

$$z = \exp(-\pi\varepsilon_1 - i\theta_2) + \exp(-\pi\varepsilon_2 - i\theta_1).$$

Moreover, one has

$$\varepsilon_1^{(0)} = \varepsilon_2^{(0)} = \frac{e}{2\pi} := \varepsilon^{(0)}. \tag{36}$$

It remains to compute the quantities θ_j (up to $O(k^{-1})$). The principal actions are easily computed:

$$c_0(\gamma_1) = -\pi, \quad c_0(\gamma_2) = 3\pi, \quad c_0(\gamma_1) = -2\pi. \tag{37}$$

Furthermore, one can check that the subprincipal actions are given by

$$\tilde{c}_1(\gamma_1) = 2\varepsilon^{(0)} \ln \frac{8}{\pi} = \tilde{c}_1(\gamma_2), \quad \tilde{c}_1(\gamma_3) = 0. \tag{38}$$

Finally, one has

$$\tilde{\epsilon}(\gamma_1) = \tilde{\epsilon}(\gamma_2) = \tilde{\epsilon}(\gamma_3) = 0. \tag{39}$$

The results thus obtained are displayed in Figure 15.

Acknowledgements

References

[Bargmann 1961] V. Bargmann, "On a Hilbert space of analytic functions and an associated integral transform", *Comm. Pure Appl. Math.* **14** (1961), 187–214. MR 28 #486 Zbl 0107.09102

[Bargmann 1967] V. Bargmann, "On a Hilbert space of analytic functions and an associated integral transform, II: A family of related function spaces. Application to distribution theory", *Comm. Pure Appl. Math.* **20** (1967), 1–101. MR 34 #1836 Zbl 0149.09601

[Berge 1973] C. Berge, *Graphs and hypergraphs*, North-Holland Mathematical Library **6**, North-Holland, New York, 1973. MR 50 #9640 Zbl 0254.05101

[Bloch et al. 2003] A. Bloch, F. Golse, T. Paul, and A. Uribe, "Dispersionless Toda and Toeplitz operators", *Duke Math. J.* **117**:1 (2003), 157–196. MR 2003m:37121 Zbl 1024.37047

[Bolsinov and Fomenko 2004] A. V. Bolsinov and A. T. Fomenko, *Integrable Hamiltonian systems: Geometry, topology, classification*, Chapman & Hall/CRC, Boca Raton, FL, 2004. MR 2004j:37106 Zbl 1056.37075

[Borthwick et al. 1998] D. Borthwick, T. Paul, and A. Uribe, "Semiclassical spectral estimates for Toeplitz operators", *Ann. Inst. Fourier (Grenoble)* **48**:4 (1998), 1189–1229. MR 2000c:58048 Zbl 0920.58059

[Charles 2003a] L. Charles, "Berezin–Toeplitz operators, a semi-classical approach", *Comm. Math. Phys.* **239**:1-2 (2003), 1–28. MR 2004m:53158 Zbl 1059.47030

[Charles 2003b] L. Charles, "Quasimodes and Bohr–Sommerfeld conditions for the Toeplitz operators", *Comm. Partial Differential Equations* **28**:9-10 (2003), 1527–1566. MR 2005d:53142 Zbl 1038.53086

[Charles 2006] L. Charles, "Symbolic calculus for Toeplitz operators with half-form", *J. Symplectic Geom.* **4**:2 (2006), 171–198. MR 2007k:58042 Zbl 1123.58016

[Charles and Marché 2011] L. Charles and J. Marché, "Knot state asymptotics, I: AJ conjecture and abelian representations", preprint, 2011, Available at http://www.math.jussieu.fr/~marche/knot1.pdf.

[Charles and Vũ Ngọc 2008] L. Charles and S. Vũ Ngọc, "Spectral asymptotics via the semiclassical Birkhoff normal form", *Duke Math. J.* **143**:3 (2008), 463–511. MR 2009d:58049 Zbl 1154.58015

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, 1999. MR 2001b:35237 Zbl 0926.35002

[Guillemin and Schaeffer 1977] V. Guillemin and D. Schaeffer, "On a certain class of Fuchsian partial differential equations", *Duke Math. J.* **44**:1 (1977), 157–199. MR 55 #3504 Zbl 0356.35080

[Helffer and Sjöstrand 1988] B. Helffer and J. Sjöstrand, *Analyse semi-classique pour l'équation de Harper (avec application à l'équation de Schrödinger avec champ magnétique)*, Mém. Soc. Math. France (N.S.) **34**, Société Mathématique de France, Paris, 1988. MR 91d:81024 Zbl 0714.34130

[Le Floch 2014] Y. Le Floch, "Singular Bohr–Sommerfeld conditions for 1D Toeplitz operators: Elliptic case", *Comm. Partial Differential Equations* **39**:2 (2014), 213–243. MR 3169784 Zbl 06291152

[Ma 2010] X. Ma, "Geometric quantization on Kähler and symplectic manifolds", pp. 785–810 in *Proceedings of the International Congress of Mathematicians, II* (Hyderabad, 2010), edited by R. Bhatia et al., Hindustan Book Agency, New Delhi, 2010. MR 2012g:53188 Zbl 1229.53088

[Ma and Marinescu 2008] X. Ma and G. Marinescu, "Toeplitz operators on symplectic manifolds", *J. Geom. Anal.* **18**:2 (2008), 565–611. MR 2010f:53161 Zbl 1152.81030

[Boutet de Monvel and Guillemin 1981] L. Boutet de Monvel and V. Guillemin, *The spectral theory of Toeplitz operators*, Annals of Mathematics Studies **99**, Princeton University Press, 1981. MR 85j:58141 Zbl 0469.47021

[Nonnenmacher and Voros 1997] S. Nonnenmacher and A. Voros, "Eigenstate structures around a hyperbolic point", *J. Phys. A* **30**:1 (1997), 295–315. MR 98e:81063 Zbl 0922.58026

[Oshemkov 1994] A. A. Oshemkov, "Morse functions on two-dimensional surfaces: Coding of singularities", *Trudy Mat. Inst. Steklov.* **205**:Novye Rezult. v Teor. Topol. Klassif. Integr. Sistem (1994), 131–140. In Russian; translated in *Proc. Steklov Inst. Math.* **205** (1995), 119–127. MR 97m:57045 Zbl 0995.37007

[Robert 1987] D. Robert, *Autour de l'approximation semi-classique*, Progress in Mathematics **68**, Birkhäuser, Boston, 1987. MR 89g:81016 Zbl 0621.35001

[Schlichenmaier 2010] M. Schlichenmaier, "Berezin–Toeplitz quantization for compact Kähler manifolds: A review of results", *Adv. Math. Phys.* (2010), Art. ID 927280. MR 2011g:53201 Zbl 1207.81049

[Colin de Verdière 2009] Y. Colin de Verdière, "Méthodes semi-classiques et théorie spectrale", Notes de cours, 2009, Available at http://www-fourier.ujf-grenoble.fr/~ycolver/All-Articles/93b.pdf.

[Colin de Verdière and Parisse 1994a] Y. Colin de Verdière and B. Parisse, "Équilibre instable en régime semi-classique, I: Concentration microlocale", *Comm. Partial Differential Equations* **19**:9-10 (1994), 1535–1563. MR 96b:58112 Zbl 0819.35116

[Colin de Verdière and Parisse 1994b] Y. Colin de Verdière and B. Parisse, "Équilibre instable en régime semi-classique, II: Conditions de Bohr–Sommerfeld", *Ann. Inst. H. Poincaré Phys. Théor.* **61**:3 (1994), 347–367. MR 97a:81041 Zbl 0845.35076

[Colin de Verdière and Parisse 1999] Y. Colin de Verdière and B. Parisse, "Singular Bohr–Sommerfeld rules", *Comm. Math. Phys.* **205**:2 (1999), 459–500. MR 2000k:81088 Zbl 1157.81310

[Colin de Verdière and Vey 1979] Y. Colin de Verdière and J. Vey, "Le lemme de Morse isochore", *Topology* **18**:4 (1979), 283–293. MR 80k:57059 Zbl 0441.58003

[Colin de Verdière and Vũ Ngọc 2003] Y. Colin de Verdière and S. Vũ Ngọc, "Singular Bohr–Sommerfeld rules for 2D integrable systems", *Ann. Sci. École Norm. Sup.* (4) **36**:1 (2003), 1–55. MR 2004h:58040 Zbl 1028.81026

[Vũ Ngọc 1998] S. V. Ngọc, *Sur le spectre des systèmes complètement intégrables semi-classiques avec singularités*, Ph.D. thesis, Université Grenoble 1, 1998, Available at http://perso.univ-rennes1.fr/san.vu-ngoc/articles/newthese.pdf.

[Vũ Ngọc 2000] S. Vũ Ngọc, "Bohr–Sommerfeld conditions for integrable systems with critical manifolds of focus-focus type", *Comm. Pure Appl. Math.* **53**:2 (2000), 143–217. MR 2001m:81081 Zbl 1027.81012

[Zelditch 1997] S. Zelditch, "Index and dynamics of quantized contact transformations", *Ann. Inst. Fourier* (*Grenoble*) **47**:1 (1997), 305–363. MR 99a:58082 Zbl 0865.47018

[Zelditch 2014] S. Zelditch, "Quantum dynamics from the semi-classical viewpoint", preprint, 2014, Available at http://www.math.northwestern.edu/~zelditch/Preprints/OXUPRESSFV.dvi.

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, Amer. Math. Soc., Providence, RI, 2012. MR 2952218 Zbl 1252.58001

YOHANN LE FLOCH: ylefloch@post.tau.ac.il
School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 6997801, Israel

RESOLVENT ESTIMATES FOR THE MAGNETIC SCHRÖDINGER OPERATOR

GEORGI VODEV

We prove optimal high-frequency resolvent estimates for self-adjoint operators of the form

$$G = -\Delta + ib(x) \cdot \nabla + i\nabla \cdot b(x) + V(x)$$

on $L^2(\mathbb{R}^n)$, $n \geq 3$, where $b(x)$ and $V(x)$ are large magnetic and electric potentials, respectively.

1. Introduction and statement of results

Let Δ be the (negative) Euclidean Laplacian on $\mathbb{R}^n$. It is well-known that the self-adjoint realization G_0 of the operator $-\Delta$ on $L^2(\mathbb{R}^n)$ has an absolutely continuous spectrum consisting of the interval $[0, +\infty)$ and satisfies the resolvent estimate

$$\left\| \langle x \rangle^{-s} \partial_x^{\alpha_1} (G_0 - \lambda^2 \pm i\varepsilon)^{-1} \partial_x^{\alpha_2} \langle x \rangle^{-s} \right\|_{L^2 \to L^2} \leq C \lambda^{|\alpha_1| + |\alpha_2| - 1}, \quad \lambda \geq 1, \tag{1-1}$$

for all multi-indices α_1 and α_2 such that $|\alpha_1| + |\alpha_2| \leq 2$, where $s > \frac{1}{2}$, $0 < \varepsilon \leq 1$, and the constant $C > 0$ does not depend on λ or ε. The same estimate still holds (see [Cardoso and Vodev 2002; Rodnianski and Tao 2011], for example) for λ large enough for perturbations of the form $-\Delta + V(x)$, where V is a real-valued function satisfying the conditions below. Note that (1-1) for $\alpha_1 = \alpha_2 = 0$ together with the ellipticity of the operator G_0 imply that the estimate (1-1) holds for all multi-indices α_1 and α_2 such that $|\alpha_1| + |\alpha_2| \leq 2$. This fact remains valid for more general elliptic perturbations of $-\Delta$.

The purpose of this work is to prove an analogue of (1-1) for perturbations by large magnetic and electric potentials, extending the recent results in [Cardoso et al. 2013; 2014a] to a larger class (most probably optimal) of magnetic potentials. More precisely, we study the high-frequency behavior of the resolvent of self-adjoint operators of the form

$$G = -\Delta + ib(x) \cdot \nabla + i\nabla \cdot b(x) + V(x) \quad \text{on } L^2(\mathbb{R}^n), \ n \geq 3,$$

where $b = (b_1, \ldots, b_n) \in L^\infty(\mathbb{R}^n; \mathbb{R}^n)$ is a magnetic potential and $V \in L^\infty(\mathbb{R}^n; \mathbb{R})$ is an electric potential. Hereafter, the operator $\nabla \cdot b$ is defined by $(\nabla \cdot b)u = \nabla \cdot (bu)$. Introduce the polar coordinates $r = |x|$, $w = x/|x| \in \mathbb{S}^{n-1}$. We suppose that $b(x) = b^L(x) + b^S(x)$, $V(x) = V^L(x) + V^S(x)$ with long-range parts b^L and V^L belonging to $C^1([r_0, +\infty))$, $r_0 \gg 1$ with respect to the radial variable r and satisfying the

---

conditions

$$|V^L(rw)| \le C, \tag{1-2}$$

$$\partial_r V^L(rw) \le Cr^{-1-\delta}, \tag{1-3}$$

$$|\partial_r^k b^L(rw)| \le Cr^{-k-\delta}, \quad k = 0, 1, \tag{1-4}$$

for all $r \ge r_0$, $w \in \mathbb{S}^{n-1}$, with some constants $C, \delta > 0$. The short-range parts satisfy

$$|b^S(x)| + |V^S(x)| \le C\langle x \rangle^{-1-\delta}. \tag{1-5}$$

Note that in the case $b^L \equiv 0$, $V^L \equiv 0$ and b^S, V^S satisfying (1-5), the operator G has an absolutely continuous spectrum consisting of the interval $[0, +\infty)$ with no strictly positive eigenvalues (see [Koch and Tataru 2006]). It follows from our result below that in the more general case when the long-range parts are not identically zero the spectrum of the operator G has a similar structure in an interval of the form $[a, +\infty)$ with some constant $a > 0$. Our main result is the following:

Theorem 1.1. *Under the conditions* (1-2)–(1-5), *for every $s > \frac{1}{2}$ there exist constants $C, \lambda_0 > 0$ so that for $\lambda \ge \lambda_0$, $0 < \varepsilon \le 1$, $|\alpha_1|, |\alpha_2| \le 1$, we have the estimate*

$$\left\| \langle x \rangle^{-s} \partial_x^{\alpha_1} (G - \lambda^2 \pm i\varepsilon)^{-1} \partial_x^{\alpha_2} \langle x \rangle^{-s} \right\|_{L^2 \to L^2} \le C\lambda^{|\alpha_1|+|\alpha_2|-1}. \tag{1-6}$$

This kind of resolvent estimates plays an important role in proving uniform local energy decay, dispersive, smoothing and Strichartz estimates for solutions to the corresponding wave and Schrödinger equations (see [Cardoso et al. 2013; 2014b; Erdoğan et al. 2009], for example). In particular, it follows from the above theorem that the smoothing and Strichartz estimates for solutions to the corresponding Schrödinger equation proved in [Erdoğan et al. 2009] hold true without the continuity condition on the magnetic potential.

Theorem 1.1 is proved in [Cardoso et al. 2013] assuming additionally that $b^S(x)$ is continuous with respect to the radial variable r uniformly in w. In the case $b^L \equiv 0$, $V^L \equiv 0$ and b^S, V^S satisfying (1-5), the estimate (1-6) is proved in [Erdoğan et al. 2009] under the extra assumption that $b(x)$ is continuous in x. In fact, no continuity of the magnetic potential is needed in order to have (1-6), as shown in [Cardoso et al. 2014a]. Instead, it was supposed in [Cardoso et al. 2014a] that $\operatorname{div} b^L$ and $\operatorname{div} b^S$ exist as functions in L^∞. This assumption allows us to conclude that the perturbation (which is a first-order differential operator) sends the Sobolev space H^1 into L^2, a fact used in an essential way in [Cardoso et al. 2014a]. Thus, our goal in the present paper is to remove this technical condition on the magnetic potential. To this end, we propose a new approach inspired by the *global* Carleman estimates proved recently in [Datchev 2014] in a different context. In what follows we will describe the main points of our proof.

There are two main difficulties in proving the above theorem. The first one is that, under our assumptions, the commutator of the gradient and the magnetic potential is not an L^∞ function. Consequently, the perturbation does not send the Sobolev space H^1 into L^2. Instead, it is bounded from H^1 into H^{-1}. Secondly, the magnetic potential is large, and therefore it is hard to apply perturbation arguments similar to those used in [Cardoso et al. 2013]. Thus, to prove Theorem 1.1 we first observe that (1-6) is equivalent

to a semiclassical a priori estimate on weighted Sobolev spaces (see (2-10) below). Furthermore, we derive this a priori estimate from a semiclassical Carleman estimate on weighted Sobolev spaces (see (2-7) below) with a suitably chosen phase function independent of the semiclassical parameter. To get this Carleman estimate we first prove a semiclassical Carleman estimate on weighted Sobolev spaces for the long-range part of the operator (see Theorem 2.1 below) and we then apply a perturbation argument. Note that the estimate (2-1) is valid for any phase function $\varphi(r) \in C^2(\mathbb{R})$ whose first derivative $\varphi'(r)$ is of compact support and nonnegative. The main feature of our Carleman estimate is that it is uniform with respect to the phase function φ (that is, the constant C_1 does not depend on φ), and the weight in the right-hand side is smaller than the usual one (that is, $(\langle x \rangle^{-2s} + \varphi'(|x|))^{-1/2}$ instead of $\langle x \rangle^s$). Thus, we can make this weight small on an arbitrary compact set by choosing the phase function properly. Moreover, in the right-hand side we have the better semiclassical Sobolev H^{-1} norm instead of the L^2 one, which is crucial for the application we make here. Note also that Carleman estimates similar to (2-1) and (2-7) have recently been proved in [Datchev 2014] for operators of the form $-h^2\Delta + V(x, h)$, where V is a real-valued long-range potential which is C^1 with respect to the radial variable r. There are, however, several important differences between the Carleman estimates in [Datchev 2014] and ours. First, the phase function in [Datchev 2014] is of the form $\varphi = \varphi_1(r)/h$, where φ_1 does not depend on h and must satisfy some conditions. Thus, the Carleman estimates in [Datchev 2014] lead to the conclusion that the resolvent in that case is bounded by $e^{C/h}$, $C > 0$ being a constant. Secondly, in [Datchev 2014] the Carleman estimates are not uniform with respect to the phase function and the norm in the right-hand side is L^2 (and not H^{-1}). Finally, the operator in [Datchev 2014] does not contain a magnetic potential.

To prove Theorem 2.1 we make use of methods originating from [Cardoso and Vodev 2002]. Note that in [Cardoso and Vodev 2002] the high-frequency behavior of the resolvent of operators of the form $-\Delta_g + V$ is studied, where V is a real-valued scalar potential and Δ_g is the negative Laplace–Beltrami operator on unbounded Riemannian manifolds, such as, for example, asymptotically Euclidean and hyperbolic ones. Similar techniques have been also used in [Rodnianski and Tao 2011], where actually all ranges of frequencies are covered. In these two papers, however, no perturbations by magnetic potentials are studied.

2. Proof of Theorem 1.1

Set $h = \lambda^{-1}$, $P(h) = h^2 G$, $\tilde{b}(x, h) = hb(x)$, $\tilde{b}^L(x, h) = h\chi(|x|)b^L(x)$, $\tilde{b}^S(x, h) = \tilde{b}(x, h) - \tilde{b}^L(x, h)$, $\widetilde{V}(x, h) = h^2 V(x)$, $\widetilde{V}^L(x, h) = h^2\chi(|x|)V^L(x)$, $\widetilde{V}^S(x, h) = \widetilde{V}(x, h) - \widetilde{V}^L(x, h)$, where $\chi \in C^\infty(\mathbb{R})$, $\chi(r) = 0$ for $r \leq r_0 + 1$, $\chi(r) = 1$ for $r \geq r_0 + 2$. Throughout this paper, $H^1(\mathbb{R}^n)$ will denote the Sobolev space equipped with the semiclassical norm

$$\|u\|_{H^1}^2 = \sum_{0 \leq |\alpha| \leq 1} \|\mathscr{D}_x^\alpha u\|_{L^2}^2,$$

where $\mathscr{D}_x = ih\partial_x$. Furthermore, H^{-1} will denote the dual space of H^1 with respect to the scalar product $\langle \cdot, \cdot \rangle_{L^2}$ with the norm

$$\|v\|_{H^{-1}} = \sup_{0 \neq u \in H^1} \frac{|\langle u, v \rangle_{L^2}|}{\|u\|_{H^1}}.$$

Let $\rho \in C^\infty(\mathbb{R})$ be a function independent of h such that $0 \leq \rho \leq 1$ and $\rho(\sigma) = 1$ for $\sigma \leq 0$, $\rho(\sigma) = 0$ for $\sigma \geq 1$. Define the function $\varphi(r) \in C^\infty(\mathbb{R})$ as follows: $\varphi(0) = 0$ and

$$\varphi'(r) = \tau\rho(r - A),$$

where $\tau, A \geq 1$ are parameters independent of h to be fixed later on. Introduce the operator

$$P^L(h) = -h^2\Delta + ih\tilde{b}^L(x, h) \cdot \nabla + ih\nabla \cdot \tilde{b}^L(x, h) + \tilde{V}^L(x, h)$$

and set

$$P_\varphi^L(h) = e^\varphi P^L(h)e^{-\varphi},$$

$$P_\varphi(h) = e^\varphi P(h)e^{-\varphi} = P_\varphi^L(h) + ih\tilde{b}^S(x, h) \cdot \nabla + ih\nabla \cdot \tilde{b}^S(x, h) - 2ih\tilde{b}^S(x, h) \cdot \nabla\varphi + \tilde{V}^S(x, h),$$

$$\mu(x) = \sqrt{\langle x\rangle^{-2s} + \varphi'(|x|)}.$$

In this section we will show that Theorem 1.1 follows from:

Theorem 2.1. *Suppose* (1-2), (1-3), (1-4) *hold and let* $\frac{1}{2} < s < \frac{1}{2}(1 + \delta)$. *Then, for all functions* $f \in H^1(\mathbb{R}^n)$ *such that* $\langle x\rangle^s(P_\varphi^L(h) - 1 \pm i\varepsilon)f \in H^{-1}(\mathbb{R}^n)$, *we have the a priori estimate*

$$\|\langle x\rangle^{-s}f\|_{H^1} \leq \frac{C_1}{h}\|\mu^{-1}(P_\varphi^L(h) - 1 \pm i\varepsilon)f\|_{H^{-1}} + C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|f\|_{L^2} \tag{2-1}$$

for $0 < \varepsilon \leq 1$, $0 < h \leq h_0(\tau, A) \ll 1$, *with a constant* $C_1 > 0$ *independent of* $f, \varepsilon, h, \tau, A$, *and a constant* $C_2 > 0$ *independent of* f, ε, h.

Let us first see that (2-1) implies the estimate

$$\|\langle x\rangle^{-s}f\|_{H^1} \leq \frac{2C_1}{h}\|\langle x\rangle^s(P_\varphi(h) - 1 \pm i\varepsilon)f\|_{H^{-1}} + 2C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|f\|_{L^2}. \tag{2-2}$$

Using that $\mu(x) \geq \tau^{1/2}$ for $|x| \leq A$ and $\mu(x) \geq \langle x\rangle^{-s}$ for $|x| \geq A + 1$ together with the condition (1-4), we get (for $0 < s - \frac{1}{2} \ll 1$)

$$\langle x\rangle^s\mu(x)^{-1}\left(|\tilde{b}^S(x, h)| + |\tilde{V}^S(x, h)|\right) \leq Ch(\tau^{-1/2} + A^{2s-1-\delta}), \tag{2-3}$$

$$\langle x\rangle^s\mu(x)^{-1}|\tilde{b}^S(x, h)|\,|\nabla\varphi| \leq O_{\tau,A}(h). \tag{2-4}$$

By (2-3) and (2-4),

$$\|\mu^{-1}(P_\varphi(h) - P_\varphi^L(h))\langle x\rangle^s\|_{H^1 \to H^{-1}} \leq Ch(\tau^{-1/2} + A^{2s-1-\delta} + O(h)). \tag{2-5}$$

By (2-1) and (2-5),

$\|\langle x\rangle^{-s}f\|_{H^1}$

$$\leq \frac{C_1}{h}\|\mu^{-1}(P_\varphi(h) - 1 \pm i\varepsilon)f\|_{H^{-1}} + \frac{C_1}{h}\|\mu^{-1}(P_\varphi(h) - P_\varphi^L(h))f\|_{H^{-1}} + C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|f\|_{L^2}$$

$$\leq \frac{C_1}{h}\|\langle x\rangle^s(P_\varphi(h) - 1 \pm i\varepsilon)f\|_{H^{-1}} + C(\tau^{-1/2} + A^{2s-1-\delta} + O(h))\|\langle x\rangle^{-s}f\|_{H^1} + C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|f\|_{L^2}. \tag{2-6}$$

Taking now τ^{-1}, A^{-1} and h small enough, we can absorb the second term in the right-hand side of (2-6) to obtain (2-2).

Applying (2-2) with $f = e^{\varphi} g$ we obtain the Carleman estimate

$$\|\langle x \rangle^{-s} e^{\varphi} g\|_{H^1} \leq \frac{2C_1}{h} \|\langle x \rangle^s e^{\varphi} (P(h) - 1 \pm i\varepsilon) g\|_{H^{-1}} + 2C_2 \left(\frac{\varepsilon}{h} \right)^{1/2} \|e^{\varphi} g\|_{L^2}. \tag{2-7}$$

Since the function φ does not depend on h, the function e^{φ} is bounded by positive constants both from below and from above. Thus, we deduce from (2-7) the a priori estimate

$$\|\langle x \rangle^{-s} g\|_{H^1} \leq \frac{\widetilde{C}_1}{h} \|\langle x \rangle^s (P(h) - 1 \pm i\varepsilon) g\|_{H^{-1}} + \widetilde{C}_2 \left(\frac{\varepsilon}{h} \right)^{1/2} \|g\|_{L^2} \tag{2-8}$$

with constants $\widetilde{C}_1, \widetilde{C}_2 > 0$ independent of h, ε and g. On the other hand, since the operator $P(h)$ is symmetric on $L^2(\mathbb{R}^n)$, we have

$$\varepsilon \|g\|_{L^2}^2 = \mp \operatorname{Im} \langle (P(h) - 1 \pm i\varepsilon) g, g \rangle_{L^2} \leq \gamma^{-1} h^{-1} \|\langle x \rangle^s (P(h) - 1 \pm i\varepsilon) g\|_{H^{-1}}^2 + \gamma h \|\langle x \rangle^{-s} g\|_{H^1}^2 \tag{2-9}$$

for every $\gamma > 0$. Taking γ small enough, independent of h, we deduce from (2-8) and (2-9) the a priori estimate

$$\|\langle x \rangle^{-s} g\|_{H^1} \leq \frac{C}{h} \|\langle x \rangle^s (P(h) - 1 \pm i\varepsilon) g\|_{H^{-1}} \tag{2-10}$$

with a constant $C > 0$ independent of h, ε and g. It is easy to see now that (2-10) implies the resolvent estimate (1-6) for $0 < s - \frac{1}{2} \ll 1$. On the other hand, we clearly have that, if (1-6) holds for some $s_0 > \frac{1}{2}$, it holds for all $s \geq s_0$. Hence (1-6) holds for all $s > \frac{1}{2}$.

3. Proof of Theorem 2.1

We will first prove the following:

Proposition 3.1. *Under the conditions of Theorem 2.1 we have the estimate*

$$\|\langle x \rangle^{-s} f\|_{H^1} \leq \frac{C_1}{h} \|\mu^{-1} (P_{\varphi}^L(h) - 1 \pm i\varepsilon) f\|_{L^2} + C_2 \left(\frac{\varepsilon}{h} \right)^{1/2} \|f\|_{H^1} \tag{3-1}$$

for every $0 < \varepsilon \leq 1$, $0 < h \leq h_0(\tau, A) \ll 1$, with a constant $C_1 > 0$ independent of f, ε, h, τ, A, and a constant $C_2 > 0$ independent of f, ε, h.

Proof. We pass to the polar coordinates $(r, w) \in \mathbb{R}^+ \times \mathbb{S}^{n-1}$, $r = |x|$, $w = x/|x|$, and recall that $L^2(\mathbb{R}^n) \cong L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1}, r^{n-1} \, dr \, dw)$. Denote by X the Hilbert space $L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1}, dr \, dw)$. We also denote by $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ the norm and the scalar product on $L^2(\mathbb{S}^{n-1})$. We will make use of the identity

$$r^{(n-1)/2} \Delta r^{-(n-1)/2} = \partial_r^2 + \frac{\widetilde{\Delta}_w}{r^2}, \tag{3-2}$$

where $\widetilde{\Delta}_w = \Delta_w - \frac{1}{4}(n-1)(n-3)$ and Δ_w denotes the negative Laplace–Beltrami operator on $\mathbb{S}^{n-1}$. Observe also that

$$r^{(n-1)/2} \partial_{x_j} r^{-(n-1)/2} = w_j \partial_r + r^{-1} q_j(w, \partial_w), \tag{3-3}$$

where $w_j = x_j/|x|$ and q_j is a first-order differential operator on $\mathbb{S}^{n-1}$, independent of r, antisymmetric on $L^2(\mathbb{S}^{n-1})$. It is easy to see that the operators $Q_j(w, \mathscr{D}_w) = ihq_j(w, \partial_w)$ and $\Lambda_w = -h^2\widetilde{\Delta}_w \geq 0$ satisfy the estimate

$$\|Q_j(w, \mathscr{D}_w)v\| \leq C\|\Lambda_w^{1/2}v\| + Ch\|v\| \quad \text{for all } v \in H^1(\mathbb{S}^{n-1}), \tag{3-4}$$

with a constant $C > 0$ independent of h and v. Set $u = r^{(n-1)/2}f$,

$$\mathscr{P}^\pm(h) = r^{(n-1)/2}(P^L(h) - 1 \pm i\varepsilon)r^{-(n-1)/2},$$

$$\mathscr{P}_\varphi^\pm(h) = r^{(n-1)/2}(P_\varphi^L(h) - 1 \pm i\varepsilon)r^{-(n-1)/2} = e^\varphi\mathscr{P}^\pm(h)e^{-\varphi}.$$

Using (3-2) and (3-3) we can write the operator $\mathscr{P}^\pm(h)$ in the coordinates (r, w) as follows:

$$\mathscr{P}^\pm(h) = \mathscr{D}_r^2 + \frac{\Lambda_w}{r^2} - 1 \pm i\varepsilon + \widetilde{V}^L + \sum_{j=1}^n w_j\big(\tilde{b}_j^L(rw, h)\mathscr{D}_r + \mathscr{D}_r\tilde{b}_j^L(rw, h)\big)$$
$$+ r^{-1}\sum_{j=1}^n(\tilde{b}_j^L(rw, h)Q_j(w, \mathscr{D}_w) + Q_j(w, \mathscr{D}_w)\tilde{b}_j^L(rw, h)),$$

where we have put $\mathscr{D}_r = ih\partial_r$. Since the function φ depends only on the variable r, this implies

$$\mathscr{P}_\varphi^\pm(h) = \mathscr{D}_r^2 + \frac{\Lambda_w}{r^2} - 1 \pm i\varepsilon + \widetilde{V}^L + W - 2ih\varphi'\mathscr{D}_r + \sum_{j=1}^n w_j(\tilde{b}_j^L(rw, h)\mathscr{D}_r + \mathscr{D}_r\tilde{b}_j^L(rw, h))$$
$$+ r^{-1}\sum_{j=1}^n\big(\tilde{b}_j^L(rw, h)Q_j(w, \mathscr{D}_w) + Q_j(w, \mathscr{D}_w)\tilde{b}_j^L(rw, h)\big),$$

where

$$W = -h^2\varphi'(r)^2 - h^2\varphi''(r) - 2ih\varphi'\sum_{j=1}^n w_j\tilde{b}_j^L.$$

Set

$$\Phi_s(r) = \|\langle r\rangle^{-s}u(r, \cdot)\|^2 + \|\langle r\rangle^{-s}\mathscr{D}_ru(r, \cdot)\|^2 + \|\langle r\rangle^{-s}r^{-1}\Lambda_w^{1/2}u(r, \cdot)\|^2,$$

$$\Psi_s = \|\langle r\rangle^{-s}u\|_{L^2(X)}^2 + \|\langle r\rangle^{-s}\mathscr{D}_ru\|_{L^2(X)}^2 + \|\langle r\rangle^{-s}r^{-1}\Lambda_w^{1/2}u\|_{L^2(X)}^2 = \int_0^\infty \Phi_s(r)\,dr,$$

$$M^\pm(r) = \|\mathscr{P}_\varphi^\pm(h)u(r, \cdot)\|^2,$$

$$\mathscr{M}^\pm = \int_0^\infty \mu^{-2}M^\pm(r)\,dr,$$

$$N(r) = \|u(r, \cdot)\|^2 + \|\mathscr{D}_ru(r, \cdot)\|^2,$$

$$\mathscr{N} = \int_0^\infty N(r)\,dr,$$

$$E(r) = -\langle(r^{-2}\Lambda_w - 1 + \widetilde{V}^L)u(r, \cdot), u(r, \cdot)\rangle + \|\mathscr{D}_ru(r, \cdot)\|^2$$
$$- 2r^{-1}\sum_{j=1}^n \operatorname{Re}\langle\tilde{b}_j^L(rw, h)Q_j(w, \mathscr{D}_w)u(r, \cdot), u(r, \cdot)\rangle.$$

To prove (3-1) we will make use of the method of [Cardoso and Vodev 2002; Rodnianski and Tao 2011] (used there in the case when the magnetic potential is identically zero), which is based on the observation

that the first derivative of the function $E(r)$ has a nice lower bound. The situation is more complex in the presence of a nontrivial magnetic potential, but we will show in what follows that the method still works. To be more precise, observe first that, in view of (1-1), (1-3) and (3-4), we have

$$E(r) \geq -\|r^{-1}\Lambda_w^{1/2}u(r,\cdot)\|^2 + \tfrac{1}{2}\|u(r,\cdot)\|^2 + \|\mathscr{D}_r u(r,\cdot)\|^2 - O(h)\Phi_{(1+\delta)/2}(r), \tag{3-5}$$

provided h is taken small enough. Furthermore, using that $\mathrm{Im}\langle \tilde{b}_j^L \mathscr{D}_r u, \mathscr{D}_r u\rangle = 0$ and $Q_j^* = Q_j$, it is easy to check that $E(r)$ satisfies the identity — see also [Cardoso et al. 2013; 2014a], where the same identity is used in an essential way —

$$
\begin{aligned}
E'(r) := \frac{dE(r)}{dr} \\
= \frac{2}{r}\langle r^{-2}\Lambda_w u(r,\cdot), u(r,\cdot)\rangle - \left\langle \frac{\partial \widetilde{V}^L}{\partial r}u(r,\cdot), u(r,\cdot)\right\rangle \\
- 2\sum_{j=1}^{n}\mathrm{Re}\left\langle \frac{\partial(\tilde{b}_j^L(rw,h)/r)}{\partial r}Q_j(w,\mathscr{D}_w)u(r,\cdot), u(r,\cdot)\right\rangle \\
- 2\sum_{j=1}^{n}\mathrm{Re}\left\langle w_j\frac{\partial \tilde{b}_j^L(rw,h)}{\partial r}u(r,\cdot), \mathscr{D}_r u(r,\cdot)\right\rangle + 2h^{-1}\mathrm{Im}\langle \mathscr{P}_\varphi^\pm(h)u(r,\cdot), \mathscr{D}_r u(r,\cdot)\rangle \\
\mp 2\varepsilon h^{-1}\mathrm{Re}\langle u(r,\cdot), \mathscr{D}_r u(r,\cdot)\rangle + 4\langle\varphi'\mathscr{D}_r u(r,\cdot), \mathscr{D}_r u(r,\cdot)\rangle \\
- 2h^{-1}\mathrm{Im}\langle Wu(r,\cdot), \mathscr{D}_r u(r,\cdot)\rangle. \tag{3-6}
\end{aligned}
$$

In view of (1-2), (1-3), (3-4) and (3-6), we obtain the inequality

$$
\begin{aligned}
E'(r) \geq \frac{2}{r}\|r^{-1}\Lambda_w^{1/2}u(r,\cdot)\|^2 + 4\varphi'\|\mathscr{D}_r u(r,\cdot)\|^2 - 2h^{-1}\|\mathscr{P}_\varphi^\pm(h)u(r,\cdot)\|\|\mathscr{D}_r(r,\cdot)\| \\
- O(h)\Phi_{(1+\delta)/2}(r) - O(\varepsilon h^{-1})N(r). \tag{3-7}
\end{aligned}
$$

Since $\Phi_{(1+\delta)/2}(r) \leq \Phi_s(r)$ for $\frac{1}{2} < s \leq \frac{1}{2}(1+\delta)$, we obtain from (3-7)

$$
\begin{aligned}
E'(r) \geq \frac{2}{r}\|r^{-1}\Lambda_w^{1/2}u(r,\cdot)\|^2 + 4\varphi'\|\mathscr{D}_r u(r,\cdot)\|^2 - \gamma^{-1}h^{-2}\mu^{-2}M^\pm(r) \\
- \gamma\mu^2\|\mathscr{D}_r(r,\cdot)\|^2 - O(h)\Phi_s(r) - O(\varepsilon h^{-1})N(r) \\
\geq \frac{2}{r}\|r^{-1}\Lambda_w^{1/2}u(r,\cdot)\|^2 - \gamma^{-1}h^{-2}\mu^{-2}M^\pm(r) - O(h+\gamma)\Phi_s(r) - O(\varepsilon h^{-1})N(r) \tag{3-8}
\end{aligned}
$$

for every $0 < \gamma \ll 1$. By (3-5) and (3-8),

$$\langle r\rangle^{-2s}(E(r) + rE'(r)) \geq \Phi_s(r) - \gamma^{-1}h^{-2}\mu^{-2}M^\pm(r) - O(h+\gamma)\Phi_s(r) - O(\varepsilon h^{-1})N(r). \tag{3-9}$$

Integrating (3-8) from t ($t > 0$) to $+\infty$ we get

$$E(t) = -\int_t^\infty E'(r)\,dr \leq O(\gamma^{-1}h^{-2})\mathscr{M}^\pm + O(\varepsilon h^{-1})\mathscr{N} + O(h+\gamma)\Psi_s. \tag{3-10}$$

Let $\psi > 0$ be a function independent of h and such that $\int_0^\infty \psi(r)\,dr < \infty$. Multiplying (3-10) by $\psi(t)$ and integrating from 0 to $+\infty$, we get

$$\int_0^\infty \psi(r)E(r)\,dr \leq O(\gamma^{-1}h^{-2})\mathcal{M}^\pm + O(\varepsilon h^{-1})\mathcal{N} + O(h+\gamma)\Psi_s. \tag{3-11}$$

Observe now that we have the identity

$$\int_0^\infty \langle r \rangle^{-2s}(E(r)+rE'(r))\,dr = \int_0^\infty \psi(r)E(r)\,dr, \tag{3-12}$$

where $\psi(r) = 2sr\langle r \rangle^{-2s-1}$. Combining (3-9), (3-11) and (3-12) and taking γ and h small enough, we conclude

$$\Psi_s \leq O(h^{-2})\mathcal{M}^\pm + O(\varepsilon h^{-1})\mathcal{N}. \tag{3-13}$$

Clearly, (3-13) implies (3-1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We will now show that (2-1) follows from (3-1) and the following:

Lemma 3.2. *Let $\ell \in \mathbb{R}$. Then we have the estimate*

$$\|\mu^{-\ell}(P_\varphi^L(h)-i)^{-1}\mu^\ell\|_{H^{-1}\to H^1} \leq C \tag{3-14}$$

for $0 < h \leq h_0(\tau, A) \ll 1$, with a constant $C > 0$ independent of h, τ and A.

We are going to use (3-1) with $f = (P_\varphi^L(h)-i)^{-1}g$. In view of the identity

$$1 = (1-i\mp i\varepsilon)(P_\varphi^L(h)-i)^{-1} + (P_\varphi^L(h)-i)^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)$$

and Lemma 3.2, we have

$$\|\langle x \rangle^{-s}g\|_{H^1} \leq 2\|\langle x \rangle^{-s}(P_\varphi^L(h)-i)^{-1}g\|_{H^1} + \|\langle x \rangle^{-s}(P_\varphi^L(h)-i)^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^1}A$$

$$\leq \frac{2C_1}{h}\|\mu^{-1}(P_\varphi^L(h)-i)^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{L^2}$$

$$+ 2C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|(P_\varphi^L(h)-i)^{-1}g\|_{H^1} + C_3\|(P_\varphi^L(h)-i)^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^1}$$

$$\leq \frac{2C_1}{h}\|\mu^{-1}(P_\varphi^L(h)-i)^{-1}\mu\|_{H^{-1}\to L^2}\|\mu^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^{-1}}$$

$$+ 2C_2\left(\frac{\varepsilon}{h}\right)^{1/2}\|(P_\varphi^L(h)-i)^{-1}\|_{L^2\to H^1}\|g\|_{L^2}$$

$$+ C_3\|(P_\varphi^L(h)-i)^{-1}\|_{H^{-1}\to H^1}\|(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^{-1}}$$

$$\leq \frac{C_1'}{h}\|\mu^{-1}(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^{-1}} + C_2'\left(\frac{\varepsilon}{h}\right)^{1/2}\|g\|_{L^2} + C_3'\|(P_\varphi^L(h)-1\pm i\varepsilon)g\|_{H^{-1}} \tag{3-15}$$

with a constant $C_1' > 0$ independent of ε, h, τ, A and g, and constants C_2', $C_3' > 0$ independent of ε, h and g. Since the function μ is bounded on $\mathbb{R}^n$, there exists $0 < h_0(\varphi) \ll 1$ such that for $0 < h \leq h_0$ the last term in the right-hand side of (3-15) can be bounded by the first one. Thus we get (2-1) from (3-15).

4. Proof of Lemma 3.2

It is easy to see that the estimate (3-14) holds with $\ell = 0$ and $P_\varphi^L(h)$ replaced by $-h^2\Delta$. Indeed, in this case the $L^2 \to L^2$ bound is trivial, while the $H^{-1} \to H^1$ bound follows from the fact that $\|f\|_{H^s} \sim \|(1-h^2\Delta)^{s/2}f\|_{L^2}$, $s = -1, 1$. We will use this to show that (3-14) with $\ell = 0$ still holds for first-order perturbations of the form $-h^2\Delta + Q(h)$, where

$$Q(h) = \sum_{|\alpha|=1} q_\alpha^{(1)}(x,h)\mathscr{D}_x^\alpha + \sum_{|\alpha|=1} \mathscr{D}_x^\alpha q_\alpha^{(2)}(x,h) + q_0(x,h)$$

with coefficients satisfying

$$|q_\alpha^{(1)}(x,h)| + |q_\alpha^{(2)}(x,h)| + |q_0(x,h)| \leq Ch \quad \text{for all } x \in \mathbb{R}^n. \tag{4-1}$$

Clearly, (4-1) implies

$$\|Q(h)\|_{H^1 \to H^{-1}} \leq Ch. \tag{4-2}$$

By (4-2) and the resolvent identity

$$(-h^2\Delta + Q(h) - i)^{-1} = (-h^2\Delta - i)^{-1} + (-h^2\Delta - i)^{-1}Q(h)(-h^2\Delta + Q(h) - i)^{-1},$$

we get

$$\|(-h^2\Delta + Q(h) - i)^{-1}\|_{H^{-1} \to H^1}$$
$$\leq \|(-h^2\Delta - i)^{-1}\|_{H^{-1} \to H^1} + \|(-h^2\Delta - i)^{-1}\|_{H^{-1} \to H^1}\|Q(h)\|_{H^1 \to H^{-1}}\|(-h^2\Delta + Q(h) - i)^{-1}\|_{H^{-1} \to H^1}$$
$$\leq C + O(h)\|(-h^2\Delta + Q(h) - i)^{-1}\|_{H^{-1} \to H^1}. \tag{4-3}$$

Now, taking h small enough (depending on the coefficients of $Q(h)$) we can absorb the last term in the right-hand side of (4-3) and obtain the desired estimate with a constant $C > 0$ independent of $q_\alpha^{(1)}$, $q_\alpha^{(2)}$, q_0 and h.

Thus, to prove (3-14) it suffices to show that the operator $\mu^{-\ell}P_\varphi^L(h)\mu^\ell$ equals $-h^2\Delta$ plus a first-order differential operator with coefficients satisfying (4-1). To do so, observe first that $\mu^{-\ell}P_\varphi^L(h)\mu^\ell = P_\psi^L(h)$, where $\psi = \varphi - \ell\log\mu$. Furthermore, we have

$$P_\psi^L(h) = -h^2\Delta + (i\tilde{b}^L - h\nabla\psi) \cdot h\nabla + h\nabla \cdot (i\tilde{b}^L - h\nabla\psi) - h^2|\nabla\psi|^2 - 2ih\tilde{b}^L \cdot \nabla\psi + \widetilde{V}^L.$$

It is easy to see that $|\psi'(r)|$ is bounded on $\mathbb{R}$, and hence $|\nabla\psi(|x|)|$ is bounded on $\mathbb{R}^n$. This together with the assumptions on $\tilde{b}^L$ and $\widetilde{V}^L$ imply the desired properties of the coefficients of the operator $P_\psi^L(h)$. $\quad\square$

References

[Cardoso and Vodev 2002] F. Cardoso and G. Vodev, "Uniform estimates of the resolvent of the Laplace–Beltrami operator on infinite volume Riemannian manifolds. II", *Ann. Henri Poincaré* **3**:4 (2002), 673–691. MR 2003j:58054 Zbl 1021.58016

[Cardoso et al. 2013] F. Cardoso, C. Cuevas, and G. Vodev, "High frequency resolvent estimates for perturbations by large long-range magnetic potentials and applications to dispersive estimates", *Ann. Henri Poincaré* **14**:1 (2013), 95–117. MR 3022667 Zbl 1260.81067

[Cardoso et al. 2014a] F. Cardoso, C. Cuevas, and G. Vodev, "Resolvent estimates for perturbations by large magnetic potentials", *J. Math. Phys.* **55**:2 (2014), Article ID #023502. MR 3202885 Zbl 06292944

[Cardoso et al. 2014b] F. Cardoso, C. Cuevas, and G. Vodev, "Semi-classical dispersive estimates", *Math. Z.* **278**:1–2 (2014), 251–277. MR 3267578 Zbl 06357849

[Datchev 2014] K. Datchev, "Quantitative limiting absorption principle in the semiclassical limit", *Geom. Funct. Anal.* **24**:3 (2014), 740–747. MR 3213828 Zbl 06324088

[Erdoğan et al. 2009] M. B. Erdoğan, M. Goldberg, and W. Schlag, "Strichartz and smoothing estimates for Schrödinger operators with almost critical magnetic potentials in three and higher dimensions", *Forum Math.* **21**:4 (2009), 687–722. MR 2010j:35091 Zbl 1181.35208

[Koch and Tataru 2006] H. Koch and D. Tataru, "Carleman estimates and absence of embedded eigenvalues", *Comm. Math. Phys.* **267**:2 (2006), 419–449. MR 2007k:35154 Zbl 1151.35025

[Rodnianski and Tao 2011] I. Rodnianski and T. Tao, "Effective limiting absorption principles, and application", preprint, 2011. arXiv 1105.0873

GEORGI VODEV: vodev@math.univ-nantes.fr
Département de Mathématiques, UMR 6629 du CNRS, Université de Nantes, 2, rue de la Houssiniere, BP 92208, 44332 Nantes Cedex 03, France

msp

■■
■msp

LOCAL AND NONLOCAL BOUNDARY CONDITIONS FOR μ-TRANSMISSION AND FRACTIONAL ELLIPTIC PSEUDODIFFERENTIAL OPERATORS

GERD GRUBB

A classical pseudodifferential operator P on $\mathbb{R}^n$ satisfies the μ-transmission condition relative to a smooth open subset Ω when the symbol terms have a certain twisted parity on the normal to $\partial\Omega$. As shown recently by the author, this condition assures solvability of Dirichlet-type boundary problems for P in full scales of Sobolev spaces with a singularity $d^{\mu-k}$, $d(x) = \text{dist}(x, \partial\Omega)$. Examples include fractional Laplacians $(-\Delta)^a$ and complex powers of strongly elliptic PDE.

We now introduce new boundary conditions, of Neumann type, or, more generally, nonlocal type. It is also shown how problems with data on $\mathbb{R}^n \setminus \Omega$ reduce to problems supported on $\overline{\Omega}$, and how the so-called "large" solutions arise. Moreover, the results are extended to general function spaces $F_{p,q}^s$ and $B_{p,q}^s$, including Hölder–Zygmund spaces $B_{\infty,\infty}^s$. This leads to optimal Hölder estimates, e.g., for Dirichlet solutions of $(-\Delta)^a u = f \in L_\infty(\Omega)$, $u \in d^a C^a(\overline{\Omega})$ when $0 < a < 1$, $a \neq \frac{1}{2}$.

Boundary value problems for elliptic pseudodifferential operators (ψdo's) P, on a smooth subset Ω of a Riemannian manifold Ω_1, have been studied under various hypotheses through the years. There is a well-known calculus initiated by Boutet de Monvel [Boutet de Monvel 1971; Rempel and Schulze 1982; Grubb 1984; 1990; 1996; 2009; Schrohe 2001] for integer-order ψdo's with the 0-transmission property (preserving C^∞ up to the boundary), including boundary value problems for elliptic differential operators and their inverses. There are theories treating more general operators with suitable factorizations of the principal symbol, initiated by Vishik and Eskin (see, e.g., [Eskin 1981; Shargorodsky 1994; Chkadua and Duduchava 2001]). Theories for operators without the transmission property have been developed by Schulze and coauthors, see, e.g., [Rempel and Schulze 1984; Harutyunyan and Schulze 2008], and theories where the boundary is considered as a singularity of the manifold have been developed in works of Melrose and coauthors, see, e.g., [Melrose 1993; Albin and Melrose 2009].

A category of ψdo's lying between the operators handled by the Boutet de Monvel calculus and the very general categories mentioned above consists of the ψdo's with a μ-transmission property, $\mu \in \mathbb{C}$, with respect to $\partial\Omega$. Only recently, a systematic study in H_p^s Sobolev spaces was given in [Grubb 2015a], departing from a result on such operators in C^∞-spaces by Hörmander [1985, Theorem 18.2.18] (in fact developed from the lecture notes [Hörmander 1965]). This category includes fractional Laplacians $(-\Delta)^a$ and complex powers of strongly elliptic differential operators, and also more generally polyhomogeneous ψdo's with symbol $p \sim \sum_{j\in\mathbb{N}_0} p_j$ having *even parity* (that is, $p_j(x, -\xi) = (-1)^j p_j(x, \xi)$ for $j \geq 0$)

or a twisted parity involving a factor $e^{i\pi\varrho}$. The general μ-transmission operators have such a reflection property of the symbol at $\partial\Omega$ just in the normal direction; see (1-5) below. This allows regularity and solvability results not only for s in a finite interval, but for arbitrarily large s.

The fractional Laplacian and its generalizations, often formulated as singular integral operators, are currently of interest in probability theory, finance, mathematical physics and geometry.

The work [Grubb 2015a] showed the Fredholm solvability of homogeneous or nonhomogeneous Dirichlet-type problems in large scales of Sobolev spaces, for μ-transmission ψdo's. In the present paper, we introduce more general boundary conditions and find criteria for their solvability. There are the general nonlocal conditions $\gamma_0 Bu = \psi$, where B is a μ-transmission ψdo; in addition to this, local higher-order conditions such as a Neumann-type condition involving the normal derivative at $\partial\Omega$ are treated. The case of $N \times N$ systems of ψdo is briefly considered.

Moreover, we show by use of [Johnsen 1996] that the theory also works in the Besov–Triebel–Lizorkin spaces $B_{p,q}^s$ and $F_{p,q}^s$, with special attention to the spaces $B_{\infty,\infty}^s$, which coincide with Hölder spaces C^s for $s \in \mathbb{R}_+ \setminus \mathbb{N}$. In comparison with [Grubb 2015a], this allows for a sharpening of Hölder results for $(-\Delta)^a$ (and other a-transmission operators) as follows: Let $\overline{\Omega}$ be a compact subset of $\mathbb{R}^n$. For solutions $u \in e^+ L_\infty(\Omega)$ of $r^+(-\Delta)^a u = f$,

$$f \in L_\infty(\Omega) \implies u \in e^+ d(x)^a C^a(\overline{\Omega}), \text{ when } a \in \,]0, 1[, \; a \neq \tfrac{1}{2}, \tag{0-1}$$

which is optimal in the Hölder exponent. (For $a = \frac{1}{2}$, it holds with C^a replaced by $C^{a-\varepsilon}$. Also higher regularities are treated, and optimal Hölder estimates for nonhomogeneous Dirichlet and Neumann problems are likewise shown.) In a new work, Ros-Oton and Serra [2014a] have studied integral operators with homogeneous, positive, even kernel and obtained (0-1) with C^a replaced by $C^{a-\varepsilon}$; in the smooth case this is covered by the present theory. (We are concerned with linear operators; the nonlinear implications in [Ros-Oton and Serra 2014a] are not touched here.) Such operators were treated in cases without boundary by Caffarelli and Silvestre, see, e.g., [2009].

Furthermore, we show the equivalence of Dirichlet problems for u supported in $\overline{\Omega}$ with problems prescribing a value of u on the exterior $\mathbb{R}^n \setminus \Omega$, obtaining new results for the latter, which were treated recently by, for example, Felsinger, Kassmann and Voigt [Felsinger et al. 2014] and Abatangelo [2013].

For nonhomogeneous problems the solutions can be "large" at the boundary; cf. [Abatangelo 2013] and its references. We show how the solutions have a specific power singularity when the boundary data are nontrivial.

The case $a = \frac{1}{2}$ enters as a boundary integral operator in treatments of mixed boundary value problems for elliptic differential operators. The present results are applied to mixed problems in [Grubb 2015b].

Outline. In Section 1, we briefly recall the relevant definitions of operators and spaces. Section 2 presents the basic results on Dirichlet and Neumann problems for $(-\Delta)^a$, including situations with given exterior data, and derives conclusions in Hölder spaces. Section 3 explains the extension of the general results to Besov–Triebel–Lizorkin spaces, including $B_{\infty,\infty}^s$. Section 4 introduces new nonlocal boundary conditions $\gamma_0 Bu = \psi$, as well as local Neumann-type conditions; also $N \times N$ systems of ψdo are discussed. The

Appendix illustrates the theory by treating a particular constant-coefficient case, showing how the problems for $(1 - \Delta)^a$ on $\mathbb{R}^n_+$ can be solved in full detail by explicit calculations.

1. Preliminaries

The notation of [Grubb 2015a] will be used. We shall give a brief account, and refer there for further details.

Consider a Riemannian n-dimensional C^∞ manifold Ω_1 (it can be $\mathbb{R}^n$) and an embedded smooth n-dimensional manifold $\overline{\Omega}$ with boundary $\partial\Omega$ and interior Ω. For $\Omega_1 = \mathbb{R}^n$, Ω can be $\mathbb{R}^n_\pm = \{x \in \mathbb{R}^n \mid x_n \gtrless 0\}$; we will denote $(x_1, \ldots, x_{n-1})$ by x'. In the general manifold case, $\overline{\Omega}$ is taken to be compact. For $\xi \in \mathbb{R}^n$, we let $(1 + |\xi|^2)^{\frac{1}{2}} = \langle\xi\rangle$, and denote by $[\xi]$ a positive C^∞-function equal to $|\xi|$ for $|\xi| \geq 1$ and $\geq \frac{1}{2}$ for all ξ. Restriction from $\mathbb{R}^n$ to $\mathbb{R}^n_\pm$ (or from Ω_1 to Ω or $\complement\overline{\Omega}$, respectively) is denoted by $r^\pm$, extension by zero from $\mathbb{R}^n_\pm$ to $\mathbb{R}^n$ (or from Ω or $\complement\overline{\Omega}$, respectively, to Ω_1) is denoted by $e^\pm$.

A pseudodifferential operator (ψdo) P on $\mathbb{R}^n$ is defined from a symbol $p(x, \xi)$ on $\mathbb{R}^n \times \mathbb{R}^n$ by

$$Pu = p(x, D)u = \mathrm{OP}(p(x, \xi))u = (2\pi)^{-n} \int e^{ix \cdot \xi} p(x, \xi)\hat{u}\, d\xi = \mathscr{F}^{-1}_{\xi \to x}(p(x, \xi)\hat{u}(\xi)); \qquad (1\text{-}1)$$

here, $\mathscr{F}$ is the Fourier transform $(\mathscr{F}u)(\xi) = \hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-ix \cdot \xi} u(x)\, dx$. The symbol p is assumed to be such that for some $r \in \mathbb{R}$, $\partial_x^\beta \partial_\xi^\alpha p(x, \xi)$ is $O(\langle\xi\rangle^{r-|\alpha|})$ for all α, β (defining the symbol class $S^r_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$); the symbol then has order r. The definition of P is carried over to manifolds by use of local coordinates. We refer to textbooks such as [Hörmander 1985; Taylor 1981; Grubb 2009] for the rules of calculus; [Grubb 2009] moreover gives an account of the Boutet de Monvel calculus of pseudodifferential boundary problems, see also, e.g., [Grubb 1996; Schrohe 2001]. When P is a ψdo on $\mathbb{R}^n$ or Ω_1, $P_+ = r^+ P e^+$ denotes its truncation to $\mathbb{R}^n_+$ or Ω, respectively.

Let $1 < p < \infty$ (with $1/p' = 1 - 1/p$), then we define for $s \in \mathbb{R}$ the spaces

$$\begin{aligned}
H^s_p(\mathbb{R}^n) &= \{u \in \mathscr{S}'(\mathbb{R}^n) \mid \mathscr{F}^{-1}(\langle\xi\rangle^s \hat{u}) \in L_p(\mathbb{R}^n)\}, \\
\dot{H}^s_p(\overline{\mathbb{R}}^n_+) &= \{u \in H^s_p(\mathbb{R}^n) \mid \mathrm{supp}\, u \subset \overline{\mathbb{R}}^n_+\}, \\
\overline{H}^s_p(\mathbb{R}^n_+) &= \{u \in \mathscr{D}'(\mathbb{R}^n_+) \mid u = r^+ U \text{ for some } U \in H^s_p(\mathbb{R}^n)\};
\end{aligned} \qquad (1\text{-}2)$$

here, $\mathrm{supp}\, u$ denotes the support of u. For a compact subset $\overline{\Omega}$ of Ω_1, the definition extends to define $\dot{H}^s_p(\overline{\Omega})$ and $\overline{H}^s_p(\Omega)$ by use of a finite system of local coordinates. We shall in the present paper moreover work in the Triebel–Lizorkin and Besov spaces $F^s_{p,q}$ and $B^s_{p,q}$, defined for $s \in \mathbb{R}$, $0 < p, q \leq \infty$ (we take $p < \infty$ in the F-case), and the derived spaces $\dot{F}^s_{p,q}$ and $\overline{F}^s_{p,q}$, etc. Here we refer to [Triebel 1995; Johnsen 1996] for basic definitions. ([Triebel 1995] writes $\widetilde{F}$ instead of $\dot{F}$, etc.; the present notation stems from Hörmander's works.) For a Hölder space C^t, $\dot{C}^t(\overline{\Omega})$ denotes the Hölder functions on Ω_1 supported in $\overline{\Omega}$. $B^s_{p,p}$ is also denoted by B^s_p when $p < \infty$, and $F^s_{p,p} = B^s_{p,p}$, $F^s_{p,2} = H^s_p$, $H^s_2 = B^s_2$.

We shall use the conventions $\bigcup_{\varepsilon > 0} H^{s+\varepsilon}_p = H^{s+0}_p$ and $\bigcap_{\varepsilon > 0} H^{s-\varepsilon}_p = H^{s-0}_p$, applied in a similar way for the other scales of spaces.

The results hold in particular for $B^s_{\infty,\infty}$-spaces. These are interesting because $B^s_{\infty,\infty}(\mathbb{R}^n)$ equals the Hölder space $C^s(\mathbb{R}^n)$ when $s \in \mathbb{R}_+ \setminus \mathbb{N}$. (There are similar statements for derived spaces over $\mathbb{R}^n_+$ and Ω.) The spaces $B^s_{\infty,\infty}(\mathbb{R}^n)$ can be identified with the Hölder–Zygmund spaces, often denoted $\mathscr{C}^s(\mathbb{R}^n)$ when

$s > 0$. There is a nice account of these spaces in Section 8.6 of [Hörmander 1997], where they are denoted by $C_*^s(\mathbb{R}^n)$ for all $s \in \mathbb{R}$; we shall use that label below, for simplicity of notation:

$$B_{\infty,\infty}^s = C_*^s \quad \text{for all } s \in \mathbb{R}. \tag{1-3}$$

For integer values of k one has, with $C_b^k(\mathbb{R}^n)$ denoting the space of functions with bounded continuous derivatives up to order k,

$$C_b^k(\mathbb{R}^n) \subset C^{k-1,1}(\mathbb{R}^n) \subset C_*^k(\mathbb{R}^n) \subset C^{k-0}(\mathbb{R}^n) \text{ when } k \in \mathbb{N},$$
$$C_b^0(\mathbb{R}^n) \subset L_\infty(\mathbb{R}^n) \subset C_*^0(\mathbb{R}^n), \tag{1-4}$$

and similar statements for derived spaces.

A ψdo P is called classical (or polyhomogeneous) when the symbol p has an asymptotic expansion $p(x,\xi) \sim \sum_{j \in \mathbb{N}_0} p_j(x,\xi)$ with p_j homogeneous in ξ of degree $m - j$ for all j. Then P has order m. One can even allow m to be complex; then $p \in S_{1,0}^{\mathrm{Re}\,m}(\mathbb{R}^n \times \mathbb{R}^n)$, and the operator and symbol are still said to be of order m.

Here there is an additional definition: P satisfies *the μ-transmission condition* (in short, is of type μ) for some $\mu \in \mathbb{C}$ when, in local coordinates,

$$\partial_x^\beta \partial_\xi^\alpha p_j(x, -N) = e^{\pi i(m - 2\mu - j - |\alpha|)} \partial_x^\beta \partial_\xi^\alpha p_j(x, N) \tag{1-5}$$

for all $x \in \partial\Omega$, all j, α, β, where N denotes the interior normal to $\partial\Omega$ at x. The implications of the μ-transmission property were a main subject of [Grubb 2015a].

A special role in the theory is played by the *order-reducing operators*. There is a simple definition of operators $\Xi_\pm^\mu$ on $\mathbb{R}^n$:

$$\Xi_\pm^\mu = \mathrm{OP}(([\xi'] \pm i\xi_n)^\mu)$$

(or with $[\xi']$ replaced by $\langle \xi' \rangle$); they preserve support in $\overline{\mathbb{R}}_\pm^n$, respectively. Here the function $([\xi'] \pm i\xi_n)^\mu$ does not satisfy all the estimates required for the class $S^{\mathrm{Re}\,\mu}(\mathbb{R}^n \times \mathbb{R}^n)$, but the operators are useful for some purposes. There is a more refined choice $\Lambda_\pm^\mu$ (with symbol $\lambda_\pm^\mu(\xi)$) that does satisfy all the estimates, and there is a definition $\Lambda_\pm^{(\mu)}$ in the manifold situation. These operators define homeomorphisms for all $s \in \mathbb{R}$ such as

$$\Lambda_+^{(\mu)} : \dot{H}_p^s(\overline{\Omega}) \xrightarrow{\sim} \dot{H}_p^{s-\mathrm{Re}\,\mu}(\overline{\Omega}),$$
$$\Lambda_{-,+}^{(\mu)} : \overline{H}_p^s(\Omega) \xrightarrow{\sim} \overline{H}_p^{s-\mathrm{Re}\,\mu}(\Omega); \tag{1-6}$$

here, $\Lambda_{-,+}^{(\mu)}$ is short for $r^+ \Lambda_-^{(\mu)} e^+$, suitably extended to large negative s (see Remark 1.1 and Theorem 1.3 in [Grubb 2015a]).

The following special spaces, introduced by Hörmander, are particularly adapted to μ-transmission operators P:

$$H_p^{\mu(s)}(\overline{\mathbb{R}}_+^n) = \Xi_+^{-\mu} e^+ \overline{H}_p^{s-\mathrm{Re}\,\mu}(\mathbb{R}_+^n), \quad s > \mathrm{Re}\,\mu - 1/p',$$
$$H_p^{\mu(s)}(\overline{\Omega}) = \Lambda_+^{(-\mu)} e^+ \overline{H}_p^{s-\mathrm{Re}\,\mu}(\Omega), \quad s > \mathrm{Re}\,\mu - 1/p', \tag{1-7}$$
$$\mathscr{E}_\mu(\overline{\Omega}) = e^+\{u(x) = d(x)^\mu v(x) \mid v \in C^\infty(\overline{\Omega})\};$$

namely, r^+P (of order m) maps them into $\overline{H}_p^{s-\operatorname{Re} m}(\mathbb{R}_+^n)$, $\overline{H}_p^{s-\operatorname{Re} m}(\Omega)$ and $C^\infty(\overline{\Omega})$ respectively (see [Grubb 2015a] Sections 1.3, 2, 4), and they appear as domains of realizations of P in the elliptic case. In the third line, $\operatorname{Re} \mu > -1$ (for other μ, see [Grubb 2015a]) and $d(x)$ is a C^∞-function vanishing to order 1 at $\partial\Omega$ and positive on Ω, e.g., $d(x) = \operatorname{dist}(x, \partial\Omega)$ near $\partial\Omega$. One has that $H_p^{\mu(s)}(\overline{\Omega}) \supset \dot{H}_p^s(\overline{\Omega})$, and the distributions are locally in H_p^s on Ω, but at the boundary they in general have a singular behavior. More about that in the text below.

The order-reducing operators also operate in the Besov–Triebel–Lizorkin scales of spaces, satisfying the relevant versions of (1-6), and the definitions in (1-7) extend.

2. Three basic problems for the fractional Laplacian

As a useful introduction, we start out by giving a detailed presentation of boundary problems for the basic example of the fractional Laplacian.

Let $P_a = (-\Delta)^a$, $a > 0$, and let Ω be a bounded open subset of $\mathbb{R}^n$ with a C^∞-boundary $\partial\Omega = \Sigma$. P_a, acting as $u \mapsto \mathscr{F}^{-1}(|\xi|^{2a}\hat{u})$, is a pseudodifferential operator on $\mathbb{R}^n$ of order $2a$, and it is of type a and has factorization index a relative to Ω, as defined in [Grubb 2015a]. With terminology introduced by Hörmander in the notes [1965] and now exposed in [Grubb 2015a], we consider the following problems for P_a:

(1) *The homogeneous Dirichlet problem*:

$$\begin{cases} r^+P_a u = f & \text{on } \Omega, \\ \operatorname{supp} u \subset \overline{\Omega}. \end{cases} \tag{2-1}$$

(2) *A nonhomogeneous Dirichlet problem* (with u less regular than in (2-1)):

$$\begin{cases} r^+P_a u = f & \text{on } \Omega, \\ \operatorname{supp} u \subset \overline{\Omega}, \\ d(x)^{1-a}u = \varphi & \text{on } \Sigma. \end{cases} \tag{2-2}$$

(3) *A nonhomogeneous Neumann problem*:

$$\begin{cases} r^+P_a u = f & \text{on } \Omega, \\ \operatorname{supp} u \subset \overline{\Omega}, \\ \partial_n(d(x)^{1-a}u) = \psi & \text{on } \Sigma. \end{cases} \tag{2-3}$$

It is shown in [Grubb 2015a] that (2-1) and (2-2) have good solvability properties in suitable Sobolev spaces and Hölder spaces, and we shall include (2-3) in the study below. In the following, we derive further properties of each of the three problems.

Remark 2.1. The theorems in Sections 2A and 2B below are also valid when $(-\Delta)^a$ is replaced by a general a-transmission ψdo P of order $2a$ and with factorization index a, except that bijectivity is replaced by the Fredholm property. They also hold when $\overline{\Omega}$ is a compact subset of a manifold Ω_1. The results in Section 2C extend to such operators when they are principally like $(-\Delta)^a$.

In the Appendix of this paper we have included a treatment of $(1 - \Delta)^a$ on a half-space; it is a model case where one can obtain the solvability results directly by Fourier transformation.

2A. *The homogeneous Dirichlet problem.* From the point of view of functional analysis (as used for example in [Frank and Geisinger 2014]), it is natural to define the Dirichlet realization $P_{a,D}$ as the Friedrichs extension of the symmetric operator $P_{a,0}$ in $L_2(\Omega)$ acting like $r^+ P_a$ with domain $C_0^\infty(\Omega)$. There is an associated sesquilinear form

$$p_{a,0}(u, v) = (2\pi)^{-n} \int_{\mathbb{R}^n} |\xi|^{2a} \hat{u}(\xi) \overline{\hat{v}(\xi)} \, d\xi, \quad u, v \in C_0^\infty(\Omega). \tag{2-4}$$

Since $\left(\|u\|_{L_2}^2 + \int |\xi|^{2a} |\hat{u}|^2 \, d\xi \right)^{\frac{1}{2}}$ is a norm equivalent with $\|u\|_{H_2^a}$, the completion of $C_0^\infty(\Omega)$ in this norm is $V = \dot{H}_2^a(\overline{\Omega})$, and $p_{a,0}$ extends to a continuous nonnegative symmetric sesquilinear form on V. A standard application of the Lax–Milgram lemma (e.g., as in [Grubb 2009, Chapter 12]) gives an operator $P_{a,D}$ that is selfadjoint nonnegative in $L_2(\Omega)$ and acts like $r^+ P_a : \dot{H}_2^a(\overline{\Omega}) \to \overline{H}_2^{-a}(\Omega)$, with domain

$$D(P_{a,D}) = \left\{ u \in \dot{H}_2^a(\overline{\Omega}) \mid r^+ P_a u \in L_2(\Omega) \right\}. \tag{2-5}$$

The operator has compact resolvent, and the spectrum is a nondecreasing sequence of nonnegative eigenvalues going to infinity. As we shall document below, 0 is not an eigenvalue, so $P_{a,D}$ in fact has a positive lower bound and is invertible.

The results of [Grubb 2015a, Sections 4, 7] clarify the mapping properties and solvability properties further: For $1 < p < \infty$, $r^+ P_a$ maps continuously:

$$r^+ P_a : H_p^{a(s)}(\overline{\Omega}) \to \overline{H}_p^{s-2a}(\Omega), \text{ when } s > a - 1/p'; \tag{2-6}$$

there is the regularity result

$$u \in \dot{H}_p^{a-1/p'+0}(\overline{\Omega}), \ r^+ P_a u \in \overline{H}_p^{s-2a}(\Omega) \Longrightarrow u \in H_p^{a(s)}(\overline{\Omega}), \text{ when } s > a - 1/p', \tag{2-7}$$

and the mapping (2-6) is Fredholm. (It is even bijective, as seen below.) As an application of the results for $s = 2a$, $p = 2$, we have in particular that

$$D(P_{a,D}) = H_2^{a(2a)}(\overline{\Omega}) = \Lambda_+^{(-a)} e^+ \overline{H}_2^a(\Omega); \tag{2-8}$$

see also Example 7.2 in [Grubb 2015a]. We recall from [Grubb 2015a, Theorem 5.4] that

$$H_p^{a(s)}(\overline{\Omega}) \begin{cases} = \dot{H}_p^s(\overline{\Omega}) & \text{when } a - 1/p' < s < a + 1/p, \\ \subset \dot{H}_p^{s-0}(\overline{\Omega}) & \text{when } s = a + 1/p, \\ \subset e^+ d^a \overline{H}_p^{s-a}(\Omega) + \dot{H}_p^s(\overline{\Omega}) & \text{when } s > a + 1/p, s - a - 1/p \notin \mathbb{N}, \\ \subset e^+ d^a \overline{H}_p^{s-a}(\Omega) + \dot{H}_p^{s-0}(\overline{\Omega}) & \text{when } s - a - 1/p \in \mathbb{N}. \end{cases} \tag{2-9}$$

In [Grubb 2015a, Section 7], we used Sobolev embedding theorems to draw conclusions for Hölder spaces. Slightly sharper (often optimal) results can be obtained if we use an extension of the results of [Grubb 2015a] to the general scales of Triebel–Lizorkin and Besov spaces $F_{p,q}^s$ and $B_{p,q}^s$. The extended theory will be presented in detail below in Sections 3–4; for the moment we shall borrow some results to give powerful statements for $(-\Delta)^a$, $0 < a < 1$. We recall that the notation $B_{\infty,\infty}^s$ is simplified to C_*^s,

and that C_*^s equals C^s (the ordinary Hölder space) for $s \in \mathbb{R}_+ \setminus \mathbb{N}$; see also (1-4). Moreover, as special cases of Definition 3.1 and Theorem 3.4 below for $p = q = \infty$,

$$C_*^{\mu(s)}(\overline{\Omega}) = \Lambda_+^{(-\mu)} e^+ \overline{C}_*^{s - \operatorname{Re}\mu}(\Omega) \qquad \text{for } s > \operatorname{Re}\mu - 1,$$

$$C_*^{\mu(s)}(\overline{\Omega}) \subset \begin{cases} d(x)^\mu e^+ \overline{C}_*^{s - \operatorname{Re}\mu}(\Omega) + \dot{C}_*^s(\overline{\Omega}) & \text{when } s > \operatorname{Re}\mu, \ s - \operatorname{Re}\mu \notin \mathbb{N}, \\ d(x)^\mu e^+ \overline{C}_*^{s - \operatorname{Re}\mu}(\Omega) + \dot{C}_*^{s-0}(\overline{\Omega}) & \text{when } s > \operatorname{Re}\mu, \ s - \operatorname{Re}\mu \in \mathbb{N}. \end{cases} \qquad (2\text{-}10)$$

Note also that the distributions in $C_*^{\mu(s)}(\overline{\Omega})$ are locally in C_*^s on Ω, by the ellipticity of $\Lambda_+^{(-\mu)}$.

We focus in the following on the case $0 < a < 1$, assumed in the rest of this chapter. Here we find the following results, with conclusions formulated in ordinary Hölder spaces:

Theorem 2.2. *Let $s > a - 1$. If $u \in \dot{C}_*^{a-1+\varepsilon}(\overline{\Omega})$ for some $\varepsilon > 0$ (e.g., if $u \in e^+ L_\infty(\Omega)$), and $r^+ P u \in \overline{C}_*^{s-2a}(\Omega)$, then $u \in C_*^{a(s)}(\overline{\Omega})$. The mapping $r^+ P_a$ defines a bijection*

$$r^+ P_a : C_*^{a(s)}(\overline{\Omega}) \to \overline{C}_*^{s-2a}(\Omega). \qquad (2\text{-}11)$$

In particular, for any $f \in L_\infty(\Omega)$, there exists a unique solution u of (2-1) in $C_*^{a(2a)}$; it satisfies

$$u \in e^+ d(x)^a C^a(\overline{\Omega}) \cap C^{2a}(\Omega), \ \text{when } a \neq \tfrac{1}{2},$$

$$u \in \big(e^+ d(x)^{\frac{1}{2}} C^{\frac{1}{2}}(\overline{\Omega}) + \dot{C}^{1-0}(\overline{\Omega})\big) \cap C^{1-0}(\Omega) \subset e^+ d(x)^{\frac{1}{2}} C^{\frac{1}{2}-0}(\overline{\Omega}) \cap C^{1-0}(\Omega), \ \text{when } a = \tfrac{1}{2}. \qquad (2\text{-}12)$$

For $f \in C^t(\overline{\Omega}), t > 0$, the solution satisfies

$$u \in \begin{cases} e^+ d(x)^a C^{a+t}(\overline{\Omega}) \cap C^{2a+t}(\Omega) & \text{when } a+t \text{ and } 2a+t \notin \mathbb{N}, \\ \big(e^+ d(x)^a C^{a+t-0}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega})\big) \cap C^{2a+t}(\Omega) & \text{when } a+t \in \mathbb{N}, \\ \big(e^+ d(x)^a C^{a+t}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega})\big) \cap C^{2a+t-0}(\Omega) & \text{when } 2a+t \in \mathbb{N}. \end{cases} \qquad (2\text{-}13)$$

Also, the mappings in (2-6) are bijections for $s > a - 1/p'$.

Proof. The first two statements are a special case of Theorem 3.2 below (see Example 3.3), except that we have replaced the Fredholm property with bijectivity. According to [Ros-Oton and Serra 2014b, Proposition 1.1] a weak solution (a solution in $\dot{H}_2^a(\overline{\Omega})$) of the problem (2-1) with $f \in L_\infty(\Omega)$ satisfies $\|u\|_{C^a} \leq C\|f\|_{L_\infty}$; in particular, it is unique. For $f \in \overline{H}_2^{-a}(\Omega)$, the Fredholm property of $r^+ P_a$ from $H_2^{a(a)}(\overline{\Omega}) = \dot{H}_2^a(\overline{\Omega})$ to $\overline{H}_2^{-a}(\Omega)$ is covered by [Grubb 2015a, Theorem 7.1] with $s = a$, $p = 2$. Moreover, the kernel $\mathcal{N}$ is in $\mathscr{E}_a(\overline{\Omega})$ by Theorem 3.5 below. If the kernel were nonzero, there would exist nontrivial null-solutions $u \in \mathscr{E}_a(\overline{\Omega})$, contradicting the uniqueness for $f \in L_\infty(\Omega)$ mentioned above. Thus $\mathcal{N} = 0$. Then the kernel of the Dirichlet realization $P_{a,D}$ in $L_2(\Omega)$ recalled above is likewise 0, and, since it is a selfadjoint operator with compact resolvent, it must be bijective. So the cokernel in $L_2(\Omega)$ is likewise 0. This shows the bijectivity of (2-6) in the case $s = 2a$, $p = 2$. In view of Theorem 3.5 below, this bijectivity carries over to all the other versions, including (2-6) for general $s > a - 1/p'$, and the mapping (2-11) in C_*^s-spaces for $s > a - 1$.

For (2-12) we use Theorem 3.4 (as recalled in (2-10)), noting that $\overline{C}_*^a(\Omega) = C^a(\overline{\Omega})$, that $\dot{C}_*^{2a}(\overline{\Omega}) = \dot{C}^{2a}(\overline{\Omega}) \subset d(x)^a C^a(\overline{\Omega})$ when $a \neq \tfrac{1}{2}$, and that $u \in C^{2a}(\Omega)$ by interior regularity when $a \neq \tfrac{1}{2}$, with slightly weaker statements when $a = \tfrac{1}{2}$. The rest of the statements follow similarly by use of (2-10) with $\mu = a$ and the various information on the relation between the C_*^s-spaces and standard Hölder spaces. $\qquad \square$

Ros-Oton and Serra [2014b] showed, under weaker smoothness hypotheses on Ω, the inclusion $u \in d^a C^\alpha(\overline{\Omega})$ for any α with $0 < \alpha < \min\{a, 1-a\}$, and improved it in [Ros-Oton and Serra 2014a] to $\alpha = a - \varepsilon$. They observe that $\alpha > a$ cannot be obtained, so $\alpha = a$, which we obtain in (2-12), is optimal.

We also have as shown in [Grubb 2015a, Theorem 4.4] that for functions u supported in $\overline{\Omega}$ (see the first inclusion in (2-7)),

$$r^+ P_a u \in C^\infty(\overline{\Omega}) \iff u \in \mathscr{E}_a(\overline{\Omega}) \equiv \left\{ u = e^+ d(x)^a v(x) \mid v \in C^\infty(\overline{\Omega}) \right\}. \qquad (2\text{-}14)$$

It is worth emphasizing that the functions in $\mathscr{E}_a$ have a nontrivially singular behavior at Σ when $a \notin \mathbb{N}_0$; $e^+ C^\infty(\overline{\Omega})$ and $\mathscr{E}_a(\overline{\Omega})$ are very different spaces. The appearance of a factor d^{μ_0}, where μ_0 is the factorization index, was observed in C^∞-situations also in [Eskin 1981, p. 311] and in [Chkadua and Duduchava 2001, Theorem 2.1].

The solution operator is denoted by R; its form as a composition of pseudodifferential factors was given in [Grubb 2015a].

There is another point of view on the Dirichlet problem for P_a that we shall also discuss. In a number of papers (see, e.g., [Hoh and Jacob 1996; Felsinger et al. 2014] and their references), the Dirichlet problem for P_a (and other related operators) is formulated as

$$\begin{cases} P_a U = f & \text{in } \Omega, \\ U = g & \text{on } \complement\Omega. \end{cases} \qquad (2\text{-}15)$$

Although the main aim is to determine U on Ω, the prescription of the values of U on $\complement\Omega$ is explained as necessitated by the nonlocality of P_a. As observed explicitly in [Hoh and Jacob 1996], the transmission property of [Boutet de Monvel 1971] is not satisfied; hence that theory of boundary problems for pseudodifferential operators is of no help. But now that we have the μ-transmission calculus, it is worth investigating what the methods can give.

The case $g = 0$ corresponds to the formulation (2-1). But also, in general, (2-15) can be reduced to (2-1) when the spaces are suitably chosen. For (2-15), let f be given in $\overline{H}_p^{s-2a}(\Omega)$ (with $s > a - 1/p'$), and let g be given in $\overline{H}_p^s(\complement\overline{\Omega})$; then we search for U in a Sobolev space over $\mathbb{R}^n$.

Let $G = \ell g$ be an extension of g to $H_p^s(\mathbb{R}^n)$. Then $u = U - G$ must satisfy

$$\begin{cases} r^+ P_a u = f - r^+ P_a G & \text{in } \Omega, \\ \operatorname{supp} u \subset \overline{\Omega}. \end{cases} \qquad (2\text{-}16)$$

Here $P_a G \in H_{p,\text{loc}}^{s-2a}(\mathbb{R}^n)$, so $f - r^+ P_a G \in \overline{H}_p^{s-2a}(\Omega)$.

According to our analysis of (2-1), there is a unique solution $u = R(f - r^+ P_a G) \in H_p^{a(s)}(\overline{\Omega})$ of (2-16). Then (2-15) has the solution $U = u + G \in H_p^{a(s)}(\overline{\Omega}) + H_p^s(\mathbb{R}^n)$. Moreover, there is at most one solution to (2-15) in this space, for if $U_1 = u_1 + G_1$ and $U_2 = u_2 + G_2$ are two solutions, then $v = u_1 - u_2 + G_1 - G_2$ is supported in $\overline{\Omega}$, hence lies in $H_p^{a(s)}(\overline{\Omega}) + \dot{H}_p^s(\overline{\Omega}) = H_p^{a(s)}(\overline{\Omega})$ and satisfies (2-1) with $f = 0$; hence it must be 0.

This reduction allows a study of higher regularity of the solutions. The treatment in [Felsinger et al. 2014] seems primarily directed towards the regularity involved in variational formulations ($p = 2$, $s = a$)

where Vishik and Eskin's results would be applicable; moreover, [Felsinger et al. 2014] allows a less smooth boundary.

We have shown:

Theorem 2.3. *Let $s > a - 1/p'$, and let $f \in \overline{H}_p^{s-2a}(\Omega)$ and $g \in \overline{H}_p^s(\complement\overline{\Omega})$ be given. Then the problem* (2-15) *has the unique solution $U = u + G \in H_p^{a(s)}(\overline{\Omega}) + H_p^s(\mathbb{R}^n)$, where $G \in H_p^s(\mathbb{R}^n)$ is an extension of g and*

$$u = R(f - r^+ P_a G) \in H_p^{a(s)}(\overline{\Omega}); \tag{2-17}$$

here, R is the solution operator for (2-1).

Observe in particular that the solution is independent of the choice of an extension operator $\ell : g \mapsto G$.
There is an immediate corollary for solutions in Hölder spaces (as in [Grubb 2015a, Section 7]):

Corollary 2.4. *Let $p > n/a$. For $f \in L_p(\Omega)$, $g \in C^{2a+0}(\complement\overline{\Omega}) \cap \overline{H}_p^{2a}(\complement\overline{\Omega})$, the solution of* (2-15) *according to* Theorem 2.3 *satisfies*

$$U \in e^+ d^a C^{a-n/p}(\overline{\Omega}) + C^{2a+0}(\mathbb{R}^n) \cap H_p^{2a}(\mathbb{R}^n), \tag{2-18}$$

if $2a - n/p \neq 1$. If $2a - n/p$ equals 1, we need to add the space $\dot{C}^{1-0}(\overline{\Omega})$.

Proof. The intersection with $\overline{H}_p^{2a}(\complement\overline{\Omega})$ serves as a bound at ∞. We extend g to a function $G \in C^{2a+0}(\mathbb{R}^n)$; then $G \in C^{2a+0}(\mathbb{R}^n) \cap H_p^{2a}(\mathbb{R}^n)$ (since $C^{t+0} \subset H_p^t$ over bounded sets). Theorem 2.3 now gives the existence of a solution $U = u + G$, where $u \in H_p^{a(2a)}(\overline{\Omega})$. By [Grubb 2015a, Corollary 5.5] (see (2-9) above), this is contained in $d^a C^{a-n/p}(\overline{\Omega})$ when $2a - n/p \neq 1$ ($a - 1/p$ and $a - n/p$ are already noninteger). If $2a - p/n = 1$, then we have to add the space $\dot{C}^{1-0}(\overline{\Omega})$, due to the embedding $\dot{H}_p^{1+n/p}(\overline{\Omega}) \subset \dot{C}^{1-0}(\overline{\Omega})$. $\square$

Results for problems with $f \in L_\infty(\Omega)$ or Hölder spaces were obtained in [Grubb 2015a] by letting $p \to \infty$; here we shall obtain sharper results by applying the general method to the C_*^s-scale. Repeating the proof of Theorem 2.3 in this scale, we find:

Theorem 2.5. *Let $s > a - 1$, and let $f \in \overline{C}_*^{s-2a}(\Omega)$ and $g \in \overline{C}_*^s(\complement\overline{\Omega})$ be given. Then the problem* (2-15) *has the unique solution $U = u + G \in C_*^{a(s)}(\overline{\Omega}) + C_*^s(\mathbb{R}^n)$, where $G \in C_*^s(\mathbb{R}^n)$ is an extension of g and*

$$u = R(f - r^+ P_a G) \in C_*^{a(s)}(\overline{\Omega}); \tag{2-19}$$

here, R is the solution operator for (2-1).

Let us spell this out in more detail for $s = 2a$ and $s = 2a + t$ in terms of ordinary Hölder spaces. In Corollary 2.6(1), we take g to be compactly supported in $\complement\Omega$; in (2) and (3), a very general term supported away from $\overline{\Omega}$ is added (it can in particular lie in C_*^{2a+t}). Recall from (1-4) that $L_\infty \subset C_*^0$.

Corollary 2.6. (1) *For $f \in L_\infty(\Omega)$, $g \in C_{\text{comp}}^{2a}(\complement\Omega)$, the solution of* (2-15) *according to* Theorem 2.5 *satisfies*

$$U \in e^+ d^a C^a(\overline{\Omega}) \cap C^{2a}(\Omega) + C_{\text{comp}}^{2a}(\mathbb{R}^n), \tag{2-20}$$

with C^{2a} replaced by C^{1-0} if $a = \frac{1}{2}$, and the same for C_{comp}^{2a}.

(2) *Let X be any of the function spaces $F^\sigma_{p,q}(\mathbb{R}^n)$ or $B^\sigma_{p,q}(\mathbb{R}^n)$, and denote by X_{ext} the subset of elements with support disjoint from $\overline{\Omega}$. For $f \in L_\infty(\Omega)$, $g \in C^{2a}_{\text{comp}}(\complement\Omega) + X_{\text{ext}}$, there exists a solution U of (2-15) satisfying*

$$U \in e^+ d^a C^a(\overline{\Omega}) \cap C^{2a}(\Omega) + C^{2a}_{\text{comp}}(\mathbb{R}^n) + X_{\text{ext}}, \qquad (2\text{-}21)$$

with C^{2a} replaced by C^{1-0} if $a = \frac{1}{2}$, and the same for C^{2a}_{comp}.

(3) *For $f \in C^t(\overline{\Omega})$, $g \in C^{2a+t}_{\text{comp}}(\complement\Omega) + X_{\text{ext}}$, $t > 0$, the solution according to (2) satisfies*

$$U \in e^+ d^a C^{a+t}(\overline{\Omega}) \cap C^{2a+t}(\Omega) + C^{2a+t}_{\text{comp}}(\mathbb{R}^n) + X_{\text{ext}}, \qquad (2\text{-}22)$$

with C^{a+t}, C^{2a+t} and C^{2a+t}_{comp} replaced by C^{a+t-0} C^{2a+t-0} and C^{2a+t-0}_{comp}, respectively, when the exponents hit an integer.

Proof. (1) That $g \in C^{2a}_{\text{comp}}(\complement\Omega)$ means that g is in C^{2a} over the closed set $\complement\Omega$ and vanishes outside a large ball; it extends to a function $G \in C^{2a}_{\text{comp}}(\mathbb{R}^n)$. Since $C^{2a}_{\text{comp}}(\mathbb{R}^n) \subset C^{2a}_{\text{comp},*}(\mathbb{R}^n)$, the construction in Theorem 2.5 gives a solution $U = u + G$, where u is as in (2-12).

(2) The function spaces are as described, for example, in [Johnsen 1996], with $\sigma \in \mathbb{R}$, $0 < p, q \leq \infty$ ($p < \infty$ in the F-case), and ψdo's are well-defined in these spaces. We write $g = g_1 + g_2$, where $g_1 \in C^{2a}_{\text{comp}}(\complement\Omega)$ and $g_2 \in X_{\text{ext}}$. The problem (2-15) with g replaced by g_1 has a solution $u_1 + G_1$ as in (1). For the problem (2-15) with f replaced by 0 and g replaced by g_2 we take $G_2 = g_2$. Then $P_a G_2$ is C^∞ on a neighborhood of $\overline{\Omega}$ (by the pseudolocal property of pseudodifferential operators, see, e.g., [Grubb 2009, p. 177]), so the reduced problem has a solution $u_2 \in \mathcal{E}_a(\overline{\Omega})$, and the given problem then has the solution $u_2 + g_2$.

The sum of the solutions $u_1 + G_1 + u_2 + g_2$ solves (2-15) and lies in the asserted space.

(3) This is shown in a similar way, using (2-13). $\qquad\square$

Remark 2.7. Note that according to this corollary, the effect on the solution over $\overline{\Omega}$ of an exterior contribution to g supported at a distance from $\overline{\Omega}$ is only a term in $\mathcal{E}_a(\overline{\Omega})$.

2B. *A nonhomogeneous Dirichlet problem.* For the nonhomogeneous Dirichlet problem (2-2), the crucial observation that leads to its solvability is that we can identify $\mathcal{E}_{a-1}(\overline{\Omega})/\mathcal{E}_a(\overline{\Omega})$ with $C^\infty(\Sigma)$ by use of the mapping

$$\gamma_{a-1,0} : u \mapsto \Gamma(a)(d(x)^{1-a}u)|_\Sigma \equiv \Gamma(a)\gamma_0(d^{1-a}u). \qquad (2\text{-}23)$$

(The gamma-function is included for consistency in calculations of Fourier transformations and Taylor expansions.) Namely, using normal and tangential coordinates $x = y' + y_n\vec{n}(y')$ on a tubular neighborhood $U_\delta = \{y' + y_n\vec{n}(y') \mid y' \in \Sigma, |y_n| < \delta\}$ of Σ (where $\vec{n}(y')$ denotes the interior normal at y'), we have for $v \in C^\infty(\overline{\Omega})$ that

$$v(x) = v(y' + y_n\vec{n}) = v_0(y') + y_n w(x) \quad \text{on } U_\delta \cap \overline{\Omega},$$

where $v_0 \in C^\infty(\Sigma)$ is the restriction of v to Σ (also denoted $\gamma_0 v$), and w is C^∞ on $U_\delta \cap \overline{\Omega}$. Now, when $u \in \mathscr{E}_{a-1}(\overline{\Omega})$ is written as $u = e^+ \Gamma(a)^{-1} d(x)^{a-1} v$ with $v \in C^\infty(\overline{\Omega})$ and $d(x)$ taken as y_n on U_δ, then

$$u(x) = \Gamma(a)^{-1} d(x)^{a-1} v_0(y') + \Gamma(a)^{-1} d(x)^a w(x) \quad \text{on } U_\delta \cap \overline{\Omega}, \tag{2-24}$$

where $\Gamma(a)^{-1} d(x)^a w$ is a function in $\mathscr{E}_a(\overline{\Omega})$. Here, v_0 is determined uniquely from v and hence $\gamma_{a-1,0} u$ is determined uniquely from u, and the null-space of the mapping $u \mapsto \gamma_{a-1,0} u$ is $\mathscr{E}_a(\overline{\Omega})$. See also Section 5 of [Grubb 2015a]; there it is moreover shown that the mapping

$$\gamma_{a-1,0} : \mathscr{E}_{a-1}(\overline{\Omega}) \to C^\infty(\Sigma), \quad \text{with null-space } \mathscr{E}_a(\overline{\Omega}),$$

extends to a continuous surjective mapping

$$\gamma_{a-1,0} : H_p^{(a-1)(s)}(\overline{\Omega}) \to B_p^{s-a+1/p'}(\Sigma), \quad \text{with null-space } H_p^{a(s)}(\overline{\Omega}), \quad \text{for } s > a - 1/p'. \tag{2-25}$$

Now since we have the bijectivity of $r^+ P_a$ in (2-6), we can simply adjoin the mapping (2-25) and conclude the bijectivity of

$$\begin{pmatrix} r^+ P_a \\ \gamma_{a-1,0} \end{pmatrix} : H_p^{(a-1)(s)}(\overline{\Omega}) \xrightarrow{\sim} \overline{H}_p^{s-2a}(\Omega) \times B_p^{s-a+1/p'}(\Sigma). \tag{2-26}$$

This gives the unique solvability of the problem (2-2) in these spaces. There is an inverse

$$\begin{pmatrix} R & K \end{pmatrix} = \begin{pmatrix} r^+ P_a \\ \gamma_{a-1,0} \end{pmatrix}^{-1},$$

where R is the inverse of (2-6) as introduced above and K is a mapping going from Σ to $\overline{\Omega}$. (Further details in [Grubb 2015a, Section 6].)

In C_*^s-spaces, we likewise have an extension of the mapping $\gamma_{a-1,0}$:

$$\gamma_{a-1,0} : C_*^{(a-1)(s)}(\overline{\Omega}) \to C_*^{s-a+1}(\Sigma), \quad \text{with null-space } C_*^{a(s)}(\overline{\Omega}), \quad \text{for } s > a - 1. \tag{2-27}$$

Then the result is as follows (as a special case of Theorem 3.2 below), with conclusions in Hölder spaces:

Theorem 2.8. *Let $s > a - 1$. The mapping $\{r^+ P_a, \gamma_{a-1,0}\}$ defines a bijection*

$$\{r^+ P_a, \gamma_{a-1,0}\} : C_*^{(a-1)(s)}(\overline{\Omega}) \to \overline{C}_*^{s-2a}(\Omega) \times C_*^{s-a+1}(\Sigma). \tag{2-28}$$

In particular, for any $f \in L_\infty(\Omega)$, $\varphi \in C^{a+1}(\Sigma)$, there exists a unique solution u of (2-2) in $C_*^{(a-1)(2a)}(\overline{\Omega})$; it satisfies

$$u \in \begin{cases} e^+ d(x)^{a-1} C^{a+1}(\overline{\Omega}) + \dot{C}^{2a}(\overline{\Omega}) & \text{when } a \neq \frac{1}{2}, \\ e^+ d(x)^{-\frac{1}{2}} C^{\frac{3}{2}}(\overline{\Omega}) + \dot{C}^{1-0}(\overline{\Omega}) & \text{when } a = \frac{1}{2}. \end{cases} \tag{2-29}$$

For $f \in C^t(\overline{\Omega})$, $\varphi \in C^{a+1+t}(\Sigma)$, $t > 0$, the solution satisfies

$$u \in \begin{cases} e^+ d(x)^{a-1} C^{a+1+t}(\overline{\Omega}) + \dot{C}^{2a+t}(\overline{\Omega}) & \text{when } a+t \text{ and } 2a+t \notin \mathbb{N}, \\ e^+ d(x)^{a-1} C^{a+1+t-0}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega}) & \text{when } a+t \in \mathbb{N}, \\ e^+ d(x)^{a-1} C^{a+1+t}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega}) & \text{when } 2a+t \in \mathbb{N}. \end{cases} \tag{2-30}$$

Proof. The bijectivity holds in view of the bijectivity in Theorem 2.2, and (2-27). The implications (2-29) and (2-30) follow from (2-10) with $\mu = a - 1$, together with the embedding properties recalled in Section 1. Note that since $a + 1 > 2a$, there is no need to mention an intersection with $C^{2a(+t)}(\Omega)$.   □

This gives a sharpening of Theorem 7.4 in [Grubb 2015a]. We moreover recall that as shown in [Grubb 2015a, Theorem 7.1], for functions $u \in H_p^{(a-1)(s)}(\overline{\Omega})$ for some s, p with $s > a - 1/p'$,

$$f \in C^\infty(\overline{\Omega}), \ \varphi \in C^\infty(\Sigma) \iff u \in \mathscr{E}_{a-1}(\overline{\Omega}). \tag{2-31}$$

Also for the nonhomogeneous Dirichlet problem, there exist formulations where the support condition on u is replaced by a prescription of its value on $\complement\Omega$. Abatangelo [2013] considers problems of the type

$$\begin{cases} r^+ P_a U = f & \text{on } \Omega, \\ U = g & \text{on } \complement\Omega, \\ \gamma_{a-1,0} U = \varphi & \text{on } \Sigma. \end{cases} \tag{2-32}$$

(The boundary condition in [Abatangelo 2013] takes the form of the third line when Ω is a ball, but is described in a more general way for other domains.)

For (2-32), let f, g, φ be given, with

$$\{f, g, \varphi\} \in \overline{H}_p^{s-2a}(\Omega) \times \overline{H}_p^s(\complement\overline{\Omega}) \times B_p^{s-a+1/p'}(\Sigma), \quad \text{with } s > a - 1/p'. \tag{2-33}$$

Then we search for a solution U in a Sobolev space over $\mathbb{R}^n$ that allows definition of $\gamma_{a-1,0} U$.

We want to take as G an extension of g to $H_p^s(\mathbb{R}^n)$. If $s > n/p$, such that $H_p^s(\mathbb{R}^n) \subset C^0(\mathbb{R}^n)$, we have that $\gamma_{a-1,0} : G \mapsto \Gamma(a)\gamma_0(d(x)^{1-a} G)$ is well-defined and gives 0 for $G \in H_p^s(\mathbb{R}^n)$ (since $a < 1$). If $s < 1/p$, we can take G as the extension by 0 on Ω (since $\overline{H}_p^s(\complement\overline{\Omega})$ is identified with $\dot{H}_p^s(\complement\Omega)$ when $-1/p' < s < 1/p$). If $1/p \le s \le n/p$, we can also use the extension by 0 and note that the boundary value from Ω is zero, but G is only in $H_p^{1/p-0}(\mathbb{R}^n)$. Now $U_1 = U - G$ must satisfy

$$\begin{cases} r^+ P_a U_1 = f - r^+ P_a G & \text{in } \Omega, \\ \operatorname{supp} U_1 \subset \overline{\Omega}, \\ \gamma_{a-1,0} U_1 = \varphi. \end{cases} \tag{2-34}$$

We continue the analysis for $s \notin [1/p, n/p]$; when $s > 0$, this can be achieved by taking p sufficiently large.

Since $P_a G \in H_{p,\mathrm{loc}}^{s-2a}(\mathbb{R}^n)$, $f - r^+ P_a G \in \overline{H}_p^{s-2a}(\Omega)$. In this way, we have reduced the problem to the form (2-3), where we have the solution operator $(R \ K)$, see (2-26) and the following. This implies that (2-32) has the solution

$$U = R(f - r^+ P_a G) + K\varphi + G \in H_p^{a(s)}(\overline{\Omega}) + H_p^{(a-1)(s)}(\overline{\Omega}) + H_p^s(\mathbb{R}^n). \tag{2-35}$$

It is unique, since zero data give a zero solution (as we know from (2-15) in the case $\varphi = 0$). Recall that $H_p^{a(s)}(\overline{\Omega}) \subset H_p^{(a-1)(s)}(\overline{\Omega})$.

This shows the first part of the following theorem:

Theorem 2.9. (1) *Let $s > a - 1/p'$ (if $s > 0$ assume moreover that $s \notin [1/p, n/p]$), and let f, g, φ be given as in (2-33). Let $G \in H_p^s(\mathbb{R}^n)$ be an extension of g (by zero if $s < 1/p$).*

The problem (2-32) has the unique solution (2-35) in $H_p^{(a-1)(s)}(\overline{\Omega}) + H_p^s(\mathbb{R}^n)$.

(2) *Let $s > a - 1$, $s \neq 0$, and let f, g, φ be given, with*

$$\{f, g, \varphi\} \in \overline{C}_*^{s-2a}(\Omega) \times \overline{C}_*^s(\complement\overline{\Omega}) \times C_*^{s-a+1}(\Sigma). \tag{2-36}$$

Let $G \in C_*^s(\mathbb{R}^n)$ be an extension of g (by zero if $s < 0$).

The problem (2-32) has the unique solution

$$U = R(f - r^+ P_a G) + K\varphi + G \in C_*^{(a-1)(s)}(\overline{\Omega}) + C_*^s(\mathbb{R}^n). \tag{2-37}$$

Proof. (1) was shown above, and (2) is shown in an analogous way:

For $s > 0$, the extension G has boundary value $\gamma_{a-1,0}G = \Gamma(a)\gamma_0(d^{1-a}G) = 0$ since G is continuous and $1 - a > 0$, and for $s < 0$ the boundary value from Ω is 0, since G is extended by zero (using the identification of $\overline{C}_*^s(\complement\overline{\Omega})$ with $\dot{C}_*^s(\complement\Omega)$ when $-1 < s < 0$). We then apply Theorem 2.8 to $u = U - G$. $\square$

This reduction allows a study of higher regularity of the solutions. The treatment in [Abatangelo 2013] seems primarily directed towards solutions for not very smooth data. The boundary of Ω is only assumed $C^{1,1}$ there.

Remark 2.10. When $s > a + n/p$, we note that since $H_p^{a(s)}(\overline{\Omega}) \subset e^+ d(x)^a C^0(\overline{\Omega}) \subset C^0(\mathbb{R}^n)$ (see (2-9) or [Grubb 2015a, Corollary 5.5]), the solution (2-35) is the sum of a continuous function and a term $K\varphi \in H_p^{(a-1)(s)}(\overline{\Omega})$ that stems solely from the boundary value φ. To further describe $K\varphi$, consider a localized situation, where Ω is replaced by $\mathbb{R}_+^n$, $d(x)$ is replaced by x_n, and P_a is carried over to a similar operator P (of type and factorization index a). As shown in the proof of [Grubb 2015a, Theorem 6.5], the solution $K\varphi$ (in a parametrix sense) of

$$r^+ P u = 0 \text{ in } \mathbb{R}_+^n, \quad \gamma_{a-1,0} u = \varphi \text{ at } x_n = 0,$$

is of the form $K\varphi = z + w$, where

$$z = K_{a-1,0}\varphi = \Xi_+^{1-a} e^+ K_0 \varphi = e^+ c_{a-1} x_n^{a-1} K_0 \varphi, \qquad w = -R r^+ P z \in H^{a(s)}(\overline{\mathbb{R}}_+^n) \subset C^0(\mathbb{R}^n);$$

here K_0 is the standard Poisson operator sending $\varphi \in B_p^{s-a+1/p'}(\mathbb{R}^{n-1})$ into

$$K_0 \varphi = \mathscr{F}_{\xi \to x}^{-1}(\hat{\varphi}(\xi')([\xi'] + i\xi_n)^{-1}) = \mathscr{F}_{\xi' \to x'}^{-1}(\hat{\varphi}(\xi') e^{-[\xi']x_n}) \in \overline{H}_p^{s-a+1}(\mathbb{R}_+^n),$$

with $\gamma_0 K_0 \varphi = \varphi$ (see also Corollary 5.3 and the proof of Theorem 5.4 in [Grubb 2015a]). Then

$$z = e^+ c_{a-1} x_n^{a-1} K_0 \varphi \in e^+ x_n^{a-1} \overline{H}_p^{s-a+1}(\mathbb{R}_+^n) \subset e^+ x_n^{a-1} C^{s-a+1-n/p}(\overline{\mathbb{R}}_+^n),$$

with $K_0 \varphi \neq 0$ at $\{x_n = 0\}$ when $\varphi \neq 0$. For higher s, the factor $K_0 \varphi$ lies in higher-order Sobolev and Hölder spaces, but is always nontrivial at $\{x_n = 0\}$ when $\varphi \neq 0$.

When this is carried back to the manifold situation, we have that U is the sum of a term in $C^0(\mathbb{R}^n)$ and a term $e^+ d(x)^{a-1} v$, $v \in \overline{H}_p^{s-a+1}(\Omega)$, where v is nonzero at $\partial\Omega$ when $\varphi \neq 0$. Since $a < 1$, this term blows up at the boundary.

Hence the solutions are "large" at the boundary in this precise sense, consisting of a continuous function plus a term containing the factor $d(x)^{a-1}$ nontrivially. See also (2-31).

It is a theme of [Abatangelo 2013] that there exist "large" solutions of the nonhomogeneous Dirichlet problem; we here see that this is *not an exception but a rule* of the setup, provided naturally by the part of the solution mapping going from Σ to $\overline{\Omega}$.

Theorem 2.9(1) gives the following result in Hölder spaces when $f \in L_p(\Omega) = \overline{H}_p^0(\Omega)$.

Corollary 2.11. *Let* $p > n/a$. *For* $f \in L_p(\Omega)$, $g \in C^{2a+0}(\complement\Omega) \cap \overline{H}_p^{2a}(\complement\overline{\Omega})$ *and* $\varphi \in C^{a+1/p'+0}(\Sigma)$, *the solution* U *of* (2-32) *according to* Theorem 2.8 *satisfies*

$$U \in e^+ d^{a-1} C^{a+1-n/p}(\overline{\Omega}) + \dot{C}^{2a-n/p}(\overline{\Omega}) + C^{2a+0}(\mathbb{R}^n) \cap H_p^{2a}(\mathbb{R}^n), \tag{2-38}$$

with $C^{2a-n/p}$ *replaced by* C^{1-0} *if* $2a - n/p = 1$.

Proof. Note that $2a > n/p$. We extend g as in Corollary 2.4 to a function $G \in C^{2a+0}(\mathbb{R}^n) \cap H_p^{2a}(\mathbb{R}^n)$, and note that $\varphi \in C^{a+1/p'+0}(\Sigma) \subset B_p^{a+1/p'}(\Sigma)$. Theorem 2.9(1) shows that there is a (unique) solution $U = u + K\varphi + G$ with

$$u + K\varphi \in H_p^{(a-1)(2a)}(\overline{\Omega}) \subset e^+ d^{a-1} C^{a+1-n/p}(\overline{\Omega}) + \dot{C}^{2a-n/p}(\overline{\Omega})$$

(one may consult [Grubb 2015a, (7.15)]), with the mentioned modification if $2a - n/p$ is integer. $\square$

For $f \in L_\infty(\Omega)$ or $C^t(\overline{\Omega})$, we get the sharpest results by applying the statement for C_*^s-spaces:

Corollary 2.12. *(1) For* $f \in L_\infty(\Omega)$, $g \in C_{\mathrm{comp}}^{2a}(\complement\Omega)$ *and* $\varphi \in C^{a+1}(\Sigma)$, *the solution of* (2-32) *satisfies*

$$U \in e^+ d^{a-1} C^{a+1}(\overline{\Omega}) + C_{\mathrm{comp}}^{2a}(\mathbb{R}^n), \tag{2-39}$$

with C_{comp}^{2a} *replaced by* C_{comp}^{1-0} *if* $a = \frac{1}{2}$.

(2) Let X *be any of the function spaces* $F_{p,q}^\sigma(\mathbb{R}^n)$ *or* $B_{p,q}^\sigma(\mathbb{R}^n)$, *and denote by* X_{ext} *the subset of elements with support disjoint from* $\overline{\Omega}$. *For* $f \in L_\infty(\Omega)$, $g \in C_{\mathrm{comp}}^{2a}(\complement\Omega) + X_{\mathrm{ext}}$ *and* $\varphi \in C^{a+1}(\Sigma)$, *there exists a solution of* (2-32) *satisfying*

$$U \in e^+ d^{a-1} C^{a+1}(\overline{\Omega}) + C_{\mathrm{comp}}^{2a}(\mathbb{R}^n) + X_{\mathrm{ext}}, \tag{2-40}$$

with C_{comp}^{2a} *replaced by* C_{comp}^{1-0} *if* $a = \frac{1}{2}$.

(3) For $f \in C^t(\overline{\Omega})$, $g \in C_{\mathrm{comp}}^{2a+t}(\complement\Omega) + X_{\mathrm{ext}}$ *and* $\varphi \in C^{a+1+t}(\Sigma)$, *the solution according to* (2) *satisfies*

$$U \in e^+ d^{a-1} C^{a+1+t}(\overline{\Omega}) + C_{\mathrm{comp}}^{2a+t}(\mathbb{R}^n) + X_{\mathrm{ext}},$$

with C^{a+1+t} *and* C_{comp}^{2a+t} *replaced by* $C^{a+1+t-0}$ *and* $C_{\mathrm{comp}}^{2a+t-0}$, *respectively, when the exponents hit an integer.*

Proof. We apply Theorem 2.9(2) in essentially the same way as in Corollary 2.6; details can be omitted. $\square$

2C. *A nonhomogeneous Neumann problem.* The Neumann boundary value defined in connection with $(-\Delta)^a$ is

$$\gamma_{a-1,1}u = \Gamma(a+1)\gamma_0\big(\partial_n(d(x)^{1-a}u)\big); \tag{2-41}$$

it is proportional to the second coefficient in the Taylor expansion of $d^{1-a}u$ in the normal variable at the boundary (like $\gamma_0 w$ when w is as in (2-24)).

We here have, by use of Theorem 4.3 below:

Theorem 2.13. *The mapping $\{r^+P_a, \gamma_{a-1,1}\}$ defines a Fredholm operator*

$$\{r^+P_a, \gamma_{a-1,1}\} : H_p^{(a-1)(s)}(\overline{\Omega}) \to \overline{H}_p^{s-2a}(\overline{\Omega}) \times B_p^{s-a-1/p}(\Sigma), \tag{2-42}$$

for $s > a + 1/p$.

Proof. The continuity of the mapping (2-42) follows from [Grubb 2015a, Theorem 5.1] with $\mu = a - 1$, $M = 2$. The Fredholm property follows from Theorem 4.3 below in a special case (see (3-2)) by piecing together a parametrix from the parametrix construction in local coordinates given there. We use that the parametrix exists since P_a in local coordinates has principal symbol $|\xi|^{2a}$. $\qquad\square$

There is a similar version in C_*^s-spaces, with consequences for Hölder estimates:

Theorem 2.14. *Let $s > a$. The mapping $\{r^+P_a, \gamma_{a-1,1}\}$ defines a Fredholm operator*

$$\{r^+P_a, \gamma_{a-1,1}\} : C_*^{(a-1)(s)}(\overline{\Omega}) \to \overline{C}_*^{s-2a}(\Omega) \times C_*^{s-a}(\Sigma). \tag{2-43}$$

In particular, for $\{f, \psi\} \in L_\infty(\Omega) \times C^a(\Sigma)$ subject to a certain finite set of linear constraints there exists a solution u of (2-3) in $C_*^{(a-1)(2a)}(\overline{\Omega})$; it is unique modulo a finite dimensional linear subspace $\mathcal{N} \subset \mathcal{E}_{a-1}(\overline{\Omega})$ and satisfies

$$u \in \begin{cases} e^+d(x)^{a-1}C^{a+1}(\overline{\Omega}) + \dot{C}^{2a}(\overline{\Omega}) & \text{when } a \neq \frac{1}{2}, \\ e^+d(x)^{-\frac{1}{2}}C^{\frac{3}{2}}(\overline{\Omega}) + \dot{C}^{1-0}(\overline{\Omega}) & \text{when } a = \frac{1}{2}. \end{cases} \tag{2-44}$$

For $f \in C^t(\overline{\Omega})$, $\psi \in C^{a+t}(\Sigma)$, $t > 0$, the solution satisfies

$$u \in \begin{cases} e^+d(x)^{a-1}C^{a+1+t}(\overline{\Omega}) + \dot{C}^{2a+t}(\overline{\Omega}) & \text{when } a+t \text{ and } 2a+t \notin \mathbb{N}, \\ e^+d(x)^{a-1}C^{a+1+t-0}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega}) & \text{when } a+t \in \mathbb{N}, \\ e^+d(x)^{a-1}C^{a+1+t}(\overline{\Omega}) + \dot{C}^{2a+t-0}(\overline{\Omega}) & \text{when } 2a+t \in \mathbb{N}. \end{cases} \tag{2-45}$$

Proof. The first statement is the analogue of Theorem 2.13, now derived from Theorem 4.3, for $p = q = \infty$. In the next, detailed statements we formulate the Fredholm property explicitly, using also Theorem 3.5 on the smoothness of the kernel. Here the inclusions (2-44) and (2-45) follow from the description (2-10) of $C_*^{(a-1)(s)}(\overline{\Omega})$ as in the proof of Theorem 2.8. $\qquad\square$

Also in the Neumann case, one can formulate versions of the theorems with u prescribed on $\mathbb{R}^n \setminus \Omega$, and show their equivalence with the set-up for u supported in $\overline{\Omega}$; we think this is sufficiently exemplified by the treatment of the Dirichlet condition above that we can leave details to the interested reader.

3. Boundary problems in general spaces

One of the conclusions in [Grubb 2015a] of the study of the ψdo P of order $m \in \mathbb{C}$, with factorization index and type $\mu_0 \in \mathbb{C}$, was that it could be linked, by the help of the special order-reducing operators $\Lambda_\pm^{(\mu)}$, to an operator

$$Q = \Lambda_-^{(\mu_0-m)} P \Lambda_+^{(-\mu_0)} \tag{3-1}$$

of order 0 and with factorization index and type 0, which could be treated by the help of the calculus of Boutet de Monvel on H_p^s-spaces, as accounted for in [Grubb 1990]. Results for P and its boundary value problems could then be deduced from those for Q in the case of a homogeneous boundary condition. With a natural definition of boundary operators $\gamma_{\mu,k}$, nonhomogeneous boundary conditions could also be treated. In particular, we found the structure of parametrices of r^+P, with homogeneous or nonhomogeneous Dirichlet-type conditions, as compositions of operators belonging to the Boutet de Monvel calculus with the special order-reducing operators; see Theorems 4.4, 6.1 and 6.5 of [Grubb 2015a].

The results of [Grubb 1990] have been extended to the much more general families of spaces $F_{p,q}^s$ (Triebel–Lizorkin spaces) and $B_{p,q}^s$ (Besov spaces) by Johnsen [1996]. He shows that elliptic systems on a compact manifold with a smooth boundary, belonging to the Boutet de Monvel calculus, have Fredholm solvability also in these more general spaces, with C^∞ kernels and range complements (cokernels) independent of s, p, q. Here $0 < p, q \leq \infty$ is allowed for the $B_{p,q}^s$-spaces, and the same goes for the $F_{p,q}^s$-spaces, except that p is taken $< \infty$ (to avoid long explanations of exceptional cases). The parameter s is taken $> s_0$, for a suitable s_0 depending on p and the order and class of the involved operators. We refer to [Johnsen 1996] (or to Triebel's books) for detailed descriptions of the spaces, recalling just that for $1 < p < \infty$,

$$
\begin{aligned}
&F_{2,2}^s = B_{2,2}^s = H_2^s, &&L_2\text{-Sobolev spaces,} \\
&F_{p,2}^s = H_p^s, &&\text{Bessel-potential spaces,} \\
&B_{p,p}^s = B_p^s, &&\text{Besov spaces.}
\end{aligned}
\tag{3-2}
$$

Here the Bessel-potential spaces H_p^s are also called W_p^s (or $W^{s,p}$) for $s \in \mathbb{N}_0$, and the Besov spaces B_p^s are also called W_p^s (or $W^{s,p}$) for $s \in \mathbb{R}_+ \setminus \mathbb{N}$, under the common name Sobolev–Slobodetskii spaces. Recall moreover that $F_{p,p}^s = B_{p,p}^s$ for $0 < p < \infty$ (also denoted B_p^s).

We return to the general situation of $\overline{\Omega}$ smoothly embedded in a Riemannian manifold Ω_1, with $\overline{\mathbb{R}}_+^n \subset \mathbb{R}^n$ used in localizations. Hörmander's notation $\dot{F}$, $\overline{F}$ and $\dot{B}$, $\overline{B}$ will be used for the general scales, in the same way as for H_p^s; see (1-2) and the following.

In the present paper, we shall in particular be interested in the case of the scale of spaces $B_{\infty,\infty}^s = C_*^s$ (see the text around (1-3)), which gives a shortcut to sharp results on solvability in Hölder spaces.

Since we are mostly interested in results for large p, we shall assume $p \geq 1$, which simplifies the quotations from [Johnsen 1996]; namely, the condition $s > \max\{1/p-1, n/p-n\}$ simplifies to $s > 1/p-1$, since $1/p - 1 \geq n/p - n$ when $p \geq 1$. (In situations where $p < 1$ would be needed, e.g., in bootstrap regularity arguments, one can supply the presentation here with the appropriate results from [Johnsen 1996].) The usual notation $1/p' = 1 - 1/p$ is understood as 0 or 1 when $p = 1$ or ∞, respectively. We assume $p \leq \infty$ in B-cases, $p < \infty$ in F-cases, and take $0 < q \leq \infty$.

The scales $F_{p,q}^s$ and $B_{p,q}^s$ have analogous roles in definitions over $\overline{\Omega}$, but the trace mappings on them are slightly different: when $s > 1/p$,

$$\gamma_0 : \overline{F}_{p,q}^s(\Omega) \to B_{p,p}^{s-1/p}(\partial\Omega) \quad \text{and} \quad \gamma_0 : \overline{B}_{p,q}^s(\Omega) \to B_{p,q}^{s-1/p}(\partial\Omega), \tag{3-3}$$

continuously and surjectively. (One could also write $F_{p,p}^s$ instead of $B_{p,p}^s$; in [Johnsen 1996], both cases occur.)

To reduce repetitive formulations, we shall introduce the common notation

$$X_{p,q}^s \text{ stands for either } F_{p,q}^s \text{ or } B_{p,q}^s, \text{ as necessary,} \tag{3-4}$$

with the same choice in each place if the notation appears several times in the same calculation. Formulas involving boundary operators will be given explicitly in the two different cases resulting from (3-3).

In addition to the mapping and Fredholm properties established for Boutet de Monvel systems in [Johnsen 1996], we need the following generalizations of (1-6) (as in [Grubb 2015a, (1.11)–(1.20)]):

$$\begin{aligned}
\Xi_+^\mu \text{ and } \Lambda_+^\mu &: \dot{X}_{p,q}^s(\overline{\mathbb{R}}_+^n) \xrightarrow{\sim} \dot{X}_{p,q}^{s-\operatorname{Re}\mu}(\overline{\mathbb{R}}_+^n), \text{ with inverses } \Xi_+^{-\mu} \text{ and } \Lambda_+^{-\mu}, \\
\Xi_{-,+}^\mu \text{ and } \Lambda_{-,+}^\mu &: \overline{X}_{p,q}^s(\mathbb{R}_+^n) \xrightarrow{\sim} \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n), \text{ with inverses } \Xi_{-,+}^{-\mu} \text{ and } \Lambda_{-,+}^{-\mu}, \\
\Lambda_+^{(\mu)} &: \dot{X}_{p,q}^s(\overline{\Omega}) \xrightarrow{\sim} \dot{X}_{p,q}^{s-\operatorname{Re}\mu}(\overline{\Omega}), \\
\Lambda_{-,+}^{(\mu)} &: \overline{X}_{p,q}^s(\Omega) \xrightarrow{\sim} \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\Omega),
\end{aligned} \tag{3-5}$$

valid for all $s \in \mathbb{R}$. The cases with integer μ are covered by [Johnsen 1996] as a direct extension of the presentation in [Grubb 1990]; the cases of more general μ likewise extend, since the support-preserving properties extend.

We can then define (analogously to the definitions and observations in [Grubb 2015a, Sections 1.2, 1.3]):

Definition 3.1. Let $s > \operatorname{Re}\mu - 1/p'$.

(1) A distribution u on $\mathbb{R}^n$ is in $X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n)$ if and only if $\Xi_+^\mu u \in \dot{X}_{p,q}^{-1/p'+0}(\overline{\mathbb{R}}_+^n)$ and $r^+\Xi_+^\mu u \in \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n)$. In fact, $r^+\Xi_+^\mu$ maps $X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n)$ bijectively onto $\overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n)$, with inverse $\Xi_+^{-\mu}e^+$, and

$$X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n) = \Xi_+^{-\mu}e^+\overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n), \tag{3-6}$$

with the inherited norm. Here $\Lambda_+^{-\mu}$ can equivalently be used.

(2) A distribution u on Ω_1 is in $X_{p,q}^{\mu(s)}(\overline{\Omega})$ if and only if $\Lambda_+^{(\mu)}u \in \dot{X}_{p,q}^{-1/p'+0}(\overline{\Omega})$ and $r^+\Lambda_+^{(\mu)}u \in \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\Omega)$. In fact, $r^+\Lambda_+^{(\mu)}$ maps $X_{p,q}^{\mu(s)}(\overline{\Omega})$ bijectively onto $\overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\Omega)$, with inverse $\Lambda_+^{(-\mu)}e^+$, and

$$X_{p,q}^{\mu(s)}(\overline{\Omega}) = \Lambda_+^{(-\mu)}e^+\overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\Omega), \tag{3-7}$$

with the inherited norm.

The distributions in $X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n)$ and $X_{p,q}^{\mu(s)}(\overline{\Omega})$ are locally in $X_{p,q}^s$ over $\mathbb{R}_+^n$ and Ω, respectively, by interior regularity.

By use of the mapping properties of the standard trace operators γ_j described in [Johnsen 1996], and use of (3-5) above, the trace operators $\varrho_{\mu,M}$ introduced in [Grubb 2015a, Section 5] extend to the general spaces

$$\varrho_{\mu,M} = \{\gamma_{\mu,0}, \gamma_{\mu,1}, \ldots, \gamma_{\mu,M-1}\} : \begin{cases} F_{p,q}^{\mu(s)}(\overline{\Omega}) \to \prod_{0 \le j < M} B_{p,p}^{s-\operatorname{Re}\mu-j-1/p}(\partial\Omega), \\ B_{p,q}^{\mu(s)}(\overline{\Omega}) \to \prod_{0 \le j < M} B_{p,q}^{s-\operatorname{Re}\mu-j-1/p}(\partial\Omega), \end{cases} \tag{3-8}$$

for $s > \operatorname{Re}\mu + M - 1/p'$; they are surjective with kernels $F_{p,q}^{(\mu+M)(s)}(\overline{\Omega})$ and $B_{p,q}^{(\mu+M)(s)}(\overline{\Omega})$.

We can now formulate some important results from [Grubb 2015a] in these scales of spaces. Recall that when P is of type μ, it is also of type μ' for $\mu - \mu' \in \mathbb{Z}$.

Theorem 3.2. (1) *Let the ψdo P on Ω_1 be of order $m \in \mathbb{C}$ and of type $\mu \in \mathbb{C}$ relative to the boundary of the smooth compact subset $\overline{\Omega} \subset \Omega_1$. Then when $s > \operatorname{Re}\mu - 1/p'$, r^+P maps $X_{p,q}^{\mu(s)}(\overline{\Omega})$ continuously into $\overline{X}_{p,q}^{s-\operatorname{Re}m}(\Omega)$.*

(2) *Assume in addition that P is elliptic and has factorization index μ_0, where $\mu - \mu_0 \in \mathbb{Z}$. Let $s > \operatorname{Re}\mu_0 - 1/p'$. If $u \in \dot{X}_{p,q}^{\sigma}(\overline{\Omega})$ for some $\sigma > \operatorname{Re}\mu_0 - 1/p'$ and $r^+Pu \in \overline{X}_{p,q}^{s-\operatorname{Re}m}(\Omega)$, then $u \in X_{p,q}^{\mu_0(s)}(\overline{\Omega})$. The mapping r^+P defines a Fredholm operator*

$$r^+P : X_{p,q}^{\mu_0(s)}(\overline{\Omega}) \to \overline{X}_{p,q}^{s-\operatorname{Re}m}(\Omega). \tag{3-9}$$

Moreover, $\{r^+P, \gamma_{\mu_0-1,0}\}$ defines a Fredholm operator

$$\{r^+P, \gamma_{\mu_0-1,0}\} : \begin{cases} F_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega}) \to \overline{F}_{p,q}^{s-\operatorname{Re}m}(\Omega) \times B_{p,p}^{s-\operatorname{Re}\mu_0+1-1/p}(\partial\Omega), \\ B_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega}) \to \overline{B}_{p,q}^{s-\operatorname{Re}m}(\Omega) \times B_{p,q}^{s-\operatorname{Re}\mu_0+1-1/p}(\partial\Omega). \end{cases} \tag{3-10}$$

(3) *Let P be as in (2), and let $\mu' = \mu_0 - M$ for a positive integer M. Then when $s > \operatorname{Re}\mu_0 - 1/p'$, $\{r^+P, \varrho_{\mu',M}\}$ defines a Fredholm operator*

$$\{r^+P, \varrho_{\mu',M}\} : \begin{cases} F_{p,q}^{\mu'(s)}(\overline{\Omega}) \to \overline{F}_{p,q}^{s-\operatorname{Re}m}(\Omega) \times \prod_{0 \le j < M} B_{p,p}^{s-\operatorname{Re}\mu'-j-1/p}(\partial\Omega), \\ B_{p,q}^{\mu'(s)}(\overline{\Omega}) \to \overline{B}_{p,q}^{s-\operatorname{Re}m}(\Omega) \times \prod_{0 \le j < M} B_{p,q}^{s-\operatorname{Re}\mu'-j-1/p}(\partial\Omega). \end{cases} \tag{3-11}$$

Proof. (1) This is the extension of [Grubb 2015a, Theorem 4.2] to the general spaces. We recall that the proof consist of a reduction of the study of r^+P to the consideration of Q_+ (with Q as in (3-1) for $\mu = \mu_0$) of type 0; this works well in the present spaces.

(2)–(3). Here, (3-9) is obtained by a generalization of [Grubb 2015a, Theorem 4.4] and its proof to the current spaces. Now (3-11) is obtained as in [Grubb 2015a, Theorem 6.1] by adjoining the mapping (3-8) (with $\mu = \mu'$) to r^+P. Here (3-10) is the special case $M = 1$, as in [Grubb 2015a, Corollary 6.2] $\qquad\square$

The parametrices R and $\begin{pmatrix} R & K \end{pmatrix}$ described by formulas in [Grubb 2015a, Theorems 4.4, 6.5] also work in the present spaces.

Example 3.3. As an example, we have for the choice $X = B$, $p = q = \infty$, i.e., $X_{p,q}^s = B_{\infty,\infty}^s = C_*^s$, that Theorem 3.2(2) shows the following:

Let P be elliptic of order m and of type μ_0, with factorization index μ_0, and let $s > \operatorname{Re}\mu_0 - 1$. If $u \in \dot{C}_*^{\sigma}(\overline{\Omega})$ for some $\sigma > \operatorname{Re}\mu_0 - 1$ and $r^+Pu \in \overline{C}_*^{s-\operatorname{Re}m}(\Omega)$, then $u \in C_*^{\mu_0(s)}(\overline{\Omega})$. The mapping r^+P

defines a Fredholm operator

$$r^+ P : C_*^{\mu_0(s)}(\overline{\Omega}) \to \overline{C}_*^{s-\operatorname{Re} m}(\Omega). \tag{3-12}$$

Moreover, $\{r^+ P, \gamma_{\mu_0-1,0}\}$ defines a Fredholm operator

$$\{r^+ P, \gamma_{\mu_0-1,0}\} : C_*^{(\mu_0-1)(s)}(\overline{\Omega}) \to \overline{C}_*^{s-\operatorname{Re} m}(\Omega) \times C_*^{s-\operatorname{Re}\mu_0+1}(\partial\Omega). \tag{3-13}$$

For $\operatorname{Re}\mu > -1/p'$, the spaces $X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n)$ and $X_{p,q}^{\mu(s)}(\overline{\Omega})$ are further described by the following generalization of [Grubb 2015a, Theorem 5.4]:

Theorem 3.4. *One has for* $\operatorname{Re}\mu > -1$, $s > \operatorname{Re}\mu - 1/p'$, *with* $M \in \mathbb{N}$:

$$
\begin{aligned}
X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n) &\begin{cases} = \dot{X}_{p,q}^s(\overline{\mathbb{R}}_+^n) & \text{if } s - \operatorname{Re}\mu \in]-1/p', 1/p[, \\ \subset \dot{X}_{p,q}^{s-0}(\overline{\mathbb{R}}_+^n) & \text{if } s - \operatorname{Re}\mu = 1/p, \end{cases} \\
X_{p,q}^{\mu(s)}(\overline{\mathbb{R}}_+^n) &\subset e^+ x_n^\mu \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n) + \begin{cases} \dot{X}_{p,q}^s(\overline{\mathbb{R}}_+^n) & \text{if } s - \operatorname{Re}\mu \in M +]-1/p', 1/p[, \\ \dot{X}_{p,q}^{s-0}(\overline{\mathbb{R}}_+^n) & \text{if } s - \operatorname{Re}\mu = M + 1/p. \end{cases}
\end{aligned}
\tag{3-14}
$$

The inclusions (3-14) *also hold in the manifold situation, with* $\mathbb{R}_+^n$ *replaced by* Ω *and* x_n *replaced by* $d(x)$.

Proof. The first statement in (3-14) follows since $e^+ \overline{X}_{p,q}^t(\mathbb{R}_+^n) = \dot{X}_{p,q}^t(\overline{\mathbb{R}}_+^n)$ for $-1/p' < t < 1/p$; see [Johnsen 1996, (2.51)–(2.52)].

For the second statement we use the representation of u as in [Grubb 2015a, (5.13)–(5.14)], in the same way as in the proof of Theorem 5.4 there. The crucial fact is that the Poisson operator K_0 maps $\gamma_{\mu,0}u \in B_{p,p}^{s-\operatorname{Re}\mu-1/p}(\mathbb{R}^{n-1})$ and $B_{p,q}^{s-\operatorname{Re}\mu-1/p}(\mathbb{R}^{n-1})$ into $\overline{F}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n)$ and $\overline{B}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n)$, respectively (by [Johnsen 1996]), defining a term

$$v_0 = e^+ K_{\mu,0}\gamma_{\mu,0}u = c_\mu e^+ x_n^\mu K_0 \gamma_{\mu,0}u \in e^+ x_n^\mu \overline{X}_{p,q}^{s-\operatorname{Re}\mu}(\mathbb{R}_+^n),$$

with similar descriptions of terms $e^+ K_{\mu,j}\gamma_{\mu,j}u$ for j up to $M-1$, such that u by subtraction of these terms gives a term in $\dot{X}_{p,q}^s(\overline{\mathbb{R}}_+^n)$ (with s replaced by $s-0$ if $s - \operatorname{Re}\mu - 1/p$ hits an integer). $\qquad\square$

Moreover, it is important to observe the following invariance property of kernels and cokernels (typical in elliptic theory):

Theorem 3.5. *For the Fredholm operators considered in* Theorem 3.2, *the kernel is a finite-dimensional subspace* $\mathcal{N}$ *of* $\mathcal{E}_\mu(\overline{\Omega})$, *independent of the choice of* s, p, q *and* F *or* B.

There is a finite-dimensional range complement $\mathcal{M} \subset C^\infty(\overline{\Omega})$ *for* (3-9), *and* $\mathcal{M}_1 \subset C^\infty(\overline{\Omega}) \times C^\infty(\partial\Omega)^M$ *for* (3-10)–(3-11), *that is independent of the choice of* s, p, q, F, B.

Proof. This follows from the similar statement for operators in the Boutet de Monvel calculus in [Johnsen 1996, Section 5.1] when we apply the mappings $\Lambda_\pm^{(\mu)}$, etc., in the reduction of the homogeneous Dirichlet problem to a problem in the Boutet de Monvel calculus. $\qquad\square$

4. More general boundary conditions

In Theorem 3.2, we obtain the Fredholm solvability of Dirichlet-type problems defined by operators

$$\{r^+ P, \gamma_{\mu-1,0}\} : \begin{cases} F_{p,q}^{(\mu-1)(s)}(\overline{\Omega}) \to \overline{F}_{p,q}^{s-\mathrm{Re}\,m}(\Omega) \times B_{p,p}^{s-\mathrm{Re}\,\mu+1/p'}(\partial\Omega), \\ B_{p,q}^{(\mu-1)(s)}(\overline{\Omega}) \to \overline{B}_{p,q}^{s-\mathrm{Re}\,m}(\Omega) \times B_{p,q}^{s-\mathrm{Re}\,\mu+1/p'}(\partial\Omega), \end{cases} \tag{4-1}$$

for $s > \mathrm{Re}\,\mu - 1/p'$, where P is elliptic of order m, is of type μ, and has factorization index μ (called μ_0 there). In Theorem 6.5 of [Grubb 2015a] we constructed a parametrix in local coordinates, which in the Besov–Triebel–Lizorkin scales maps as

$$\left(R_D \ \ K_D \right) : \begin{cases} \overline{F}_{p,q}^{s-\mathrm{Re}\,m}(\mathbb{R}_+^n) \times B_{p,p}^{s-\mathrm{Re}\,\mu+1/p'}(\mathbb{R}^{n-1}) \to F_{p,q}^{(\mu-1)(s)}(\overline{\mathbb{R}}_+^n), \\ \overline{B}_{p,q}^{s-\mathrm{Re}\,m}(\mathbb{R}_+^n) \times B_{p,q}^{s-\mathrm{Re}\,\mu+1/p'}(\mathbb{R}^{n-1}) \to B_{p,q}^{(\mu-1)(s)}(\overline{\mathbb{R}}_+^n), \end{cases} \tag{4-2}$$

where $R_D = \Lambda_+^{-\mu} e^+ \widetilde{Q_+} \Lambda_{-,+}^{\mu-m}$ and $K_D = \Xi_+^{1-\mu} e^+ K'$ or $\Lambda_+^{1-\mu} e^+ K''$. Here $\widetilde{Q_+}$ is a parametrix of Q_+ (where Q is recalled in (3-1)), and K' and K'' are Poisson operators in the Boutet de Monvel calculus of order 0.

4A. *Boundary operators of type $\gamma_0 B$.* We shall now describe a general way to let other boundary operators enter in lieu of $\gamma_{\mu-1,0}$. The point is to reduce the problem to a problem in the Boutet de Monvel calculus (with ψdo's of type 0 and integer order). We can assume that the family of auxiliary operators $\Lambda_\pm^{(\varrho)}$ is chosen such that $(\Lambda_\pm^{(\varrho)})^{-1} = \Lambda_\pm^{(-\varrho)}$.

Theorem 4.1. *Let P be elliptic of order $m \in \mathbb{C}$ on Ω_1, having type μ and factorization index μ with respect to the smooth compact subset $\overline{\Omega}$. Let B be a ψdo of order $m_0 + \mu$ and of type μ, with m_0 integer. Consider the mapping*

$$\{r^+ P, \gamma_0 r^+ B\} : \begin{cases} F_{p,q}^{(\mu-1)(s)}(\overline{\Omega}) \to \overline{F}_{p,q}^{s-\mathrm{Re}\,m}(\Omega) \times B_{p,p}^{s-m_0-\mathrm{Re}\,\mu+1/p'}(\partial\Omega), \\ B_{p,q}^{(\mu-1)(s)}(\overline{\Omega}) \to \overline{B}_{p,q}^{s-\mathrm{Re}\,m}(\Omega) \times B_{p,q}^{s-m_0-\mathrm{Re}\,\mu+1/p'}(\partial\Omega), \end{cases} \tag{4-3}$$

for $s > \mathrm{Re}\,\mu + \max\{m_0, 0\} - 1/p'$.
(1) For $u \in X_{p,q}^{(\mu-1)(s)}(\overline{\mathbb{R}}_+^n)$, the problem

$$r^+ P u = f \text{ on } \Omega, \quad \gamma_0 r^+ B u = \psi \text{ on } \partial\Omega, \tag{4-4}$$

can be reduced to an equivalent problem

$$P'_+ w = g \text{ on } \Omega, \quad \gamma_0 B'_+ w = \psi \text{ on } \partial\Omega, \tag{4-5}$$

where $w = r^+ \Lambda_+^{(\mu-1)} u \in \overline{X}_{p,q}^{s-\mathrm{Re}\,\mu+1}(\Omega)$, $g = \Lambda_{-,+}^{(\mu-m)} f \in \overline{X}_{p,q}^{s-\mathrm{Re}\,\mu}(\Omega)$, and where

$$P' = \Lambda_-^{(\mu-m)} P \Lambda_+^{(1-\mu)}, \quad B' = B \Lambda_+^{(1-\mu)}, \tag{4-6}$$

are ψdo's of order 1 and $m_0 + 1$, respectively, and type 0.

(2) *The problem* (4-4) *is Fredholm solvable for* $s > \operatorname{Re} \mu + \max\{m_0, 0\} - 1/p'$ *if and only if the problem* (4-5) *is Fredholm solvable, as a mapping*

$$\{P'_+, \gamma_0 B'_+\} : \begin{cases} \overline{F}^{t+1}_{p,q}(\Omega) \to \overline{F}^{t}_{p,q}(\Omega) \times B^{t-m_0+1/p'}_{p,p}(\partial\Omega), \\ \overline{B}^{t+1}_{p,q}(\Omega) \to \overline{B}^{t}_{p,q}(\Omega) \times B^{t-m_0+1/p'}_{p,q}(\partial\Omega), \end{cases} \tag{4-7}$$

for $t > \max\{m_0, 0\} - 1/p'$.

(3) *The operator in* (4-7) *belongs to the Boutet de Monvel calculus; therefore Fredholm solvability holds if and only if (in addition to the invertibility of the interior symbol) the boundary symbol operator is bijective at each* $(x', \xi') \in T^*(\partial\Omega) \setminus 0$. *This can also be formulated as the unique solvability of the model problem for* (4-4) *at each* $x' \in \partial\Omega$, $\xi' \neq 0$.

(4) *In the transition between* (4-4) *and* (4-5), $\left(R'_B \ \ K'_B \right)$ *is a parametrix for* (4-5) *if and only if*

$$\left(R_B \ \ K_B \right) = \left(\Lambda^{(1-\mu)}_+ e^+ R'_B \Lambda^{(\mu-m)}_{-,+} \ \ \ \Lambda^{(1-\mu)}_+ e^+ K'_B \right) \tag{4-8}$$

is a parametrix for (4-4).

Proof. The mapping (4-3) is well-defined, since $r^+ B : X^{(\mu-1)(s)}_{p,q}(\overline{\Omega}) \to \overline{X}^{s-m_0-\operatorname{Re}\mu}_{p,q}(\Omega)$ by Theorem 3.2(1), and γ_0 acts as in (3-3).

(1) Let us go through the transition between (4-4) and (4-5), as already laid out in the formulation of the theorem.

We have from Definition 3.1 that $u \in X^{(\mu-1)(s)}_{p,q}(\overline{\Omega})$ if and only if $w = r^+ \Lambda^{(\mu-1)}_+ u \in \overline{X}^{s-\operatorname{Re}\mu+1}_{p,q}(\Omega)$; here $u = \Lambda^{(1-\mu)}_+ e^+ w$. Moreover, since $\Lambda^{(\varrho)}_{-,+} : \overline{X}^{t}_{p,q}(\Omega) \xrightarrow{\sim} \overline{X}^{t-\operatorname{Re}\varrho}_{p,q}(\Omega)$ for all ϱ and t, $f \in \overline{X}^{s-\operatorname{Re}m}_{p,q}(\Omega)$ if and only if $g = \Lambda^{(\mu-m)}_{-,+} f \in \overline{X}^{s-\operatorname{Re}\mu}_{p,q}(\Omega)$. Hence the first equation in (4-4) carries over to

$$\Lambda^{(\mu-m)}_{-,+} r^+ P \Lambda^{(1-\mu)}_+ e^+ w = g.$$

Here $\Lambda^{(\mu-m)}_{-,+} r^+ P \Lambda^{(1-\mu)}_+ e^+ w$ can be simplified to $r^+ \Lambda^{(\mu-m)}_- P \Lambda^{(1-\mu)}_+ e^+ w = P'_+ w$, as accounted for in the proof of Theorem 4.4 in [Grubb 2015a] in a similar situation. The boundary condition in (4-4) carries over to that in (4-5) since $B'_+ w = r^+ B \Lambda^{(1-\mu)}_+ e^+ w = r^+ B u$.

The order and type of the operators is clear from the definitions.

(2) Since the transition takes place by use of bijections, the Fredholm property carries over between the two situations.

(3) The model problem is the problem defined from the principal symbols of the involved operators at a boundary point x', in a local coordinate system where Ω is replaced by $\mathbb{R}^n_+$ and the operator is applied only in the x_n-direction for fixed $\xi' \neq 0$. The hereby-defined operator on $\mathbb{R}_+$ is called the boundary symbol operator in the Boutet de Monvel calculus. The first statement in (3) is just a reference to facts from the Boutet de Monvel calculus. The second statement follows immediately when the transition is applied on the principal symbol level.

(4) Finally, when $w = R'_B g + K'_B \psi$, we have

$$u = \Lambda_+^{(1-\mu)} e^+ w = \Lambda_+^{(1-\mu)} e^+ (R'_B g + K'_B \psi) = \Lambda_+^{(1-\mu)} e^+ R'_B \Lambda_{-,+}^{(\mu-m)} f + \Lambda_+^{(1-\mu)} e^+ K'_B \psi,$$

showing the last statement.   □

The search for a parametrix here requires the analysis of model problems in Sobolev-type spaces over $\mathbb{R}_+$. It can be an advantage to reduce this question to the boundary, where it suffices to investigate the ellipticity of a ψdo (i.e., invertibility of its principal symbol), as in classical treatments of differential and pseudodifferential problems.

Theorem 4.2. *Consider the problem* (4-3)–(4-4) *in* Theorem 4.1, *and its transformed version* (4-5).

(1) *The nonhomogeneous Dirichlet system for P', $\{P'_+, \gamma_0\}$, is elliptic, and has a parametrix for $s > 1/p$:*

$$\left(R'_D \ \ K'_D \right) : \begin{cases} \overline{F}_{p,q}^{s-1}(\Omega) \times B_{p,p}^{s-1/p}(\partial\Omega) \to \overline{F}_{p,q}^s(\Omega), \\ \overline{B}_{p,q}^{s-1}(\Omega) \times B_{p,q}^{s-1/p}(\partial\Omega) \to \overline{B}_{p,q}^s(\Omega). \end{cases} \tag{4-9}$$

(2) *Define*

$$S'_B = \gamma_0 B'_+ K'_D, \tag{4-10}$$

a ψdo on $\partial\Omega$ of order m_0. Then (4-3) *defines a Fredholm operator if and only if S'_B is elliptic. When this is so, if $\widetilde{S}'_B$ denotes a parametrix, then $\{r^+ P, \gamma_0 r^+ B\}$ has the parametrix $\left(R_B \ \ K_B \right)$, where*

$$R_B = \Lambda_+^{(1-\mu)}(I - K'_D \widetilde{S}'_B \gamma_0 B'_+) R'_D \Lambda_{-,+}^{(\mu-m)}, \quad K_B = \Lambda_+^{(1-\mu)} K'_D \widetilde{S}'_B. \tag{4-11}$$

Proof. We begin by discussing the solvability of the type 0 problem (4-5) with $B' = I$. Set $Q_1 = \Lambda_-^{(\mu-m)} P \Lambda_+^{(1-\mu)} \Lambda_+^{(-1)}$; it is very similar to the operator $Q = \Lambda_-^{(\mu-m)} P \Lambda_+^{(-\mu)}$ used in [Grubb 2015a, Theorems 4.2 and 4.4] being of order 0, type 0 and having factorization index 0. Then we can write

$$P' = Q_1 \Lambda_+^{(1)}, \quad P'_+ = r^+ Q_1 \Lambda_+^{(1)} e^+ = r^+ Q_1 e^+ r^+ \Lambda_+^{(1)} e^+ = Q_{1,+} \Lambda_{+,+}^{(1)}, \tag{4-12}$$

where we used that $r^- \Lambda_+^{(1)} e^+$ is 0 on $\overline{X}_{p,q}^s(\Omega)$ for $s > 1/p$.

The operator $\Lambda_+^{(1)}$ defines an elliptic (bijective) system for $s > 1/p$,

$$\{\Lambda_{+,+}^{(1)}, \gamma_0\} : \begin{cases} \overline{F}_{p,q}^s(\Omega) \xrightarrow{\sim} \overline{F}_{p,q}^{s-1}(\Omega) \times B_{p,p}^{s-1/p}(\partial\Omega), \\ \overline{B}_{p,q}^s(\Omega) \xrightarrow{\sim} \overline{B}_{p,q}^{s-1}(\Omega) \times B_{p,q}^{s-1/p}(\partial\Omega). \end{cases} \tag{4-13}$$

This is shown in [Grubb 1990, Theorem 5.1] for $q = 2$ in the F-case, and extends to the Besov–Triebel–Lizorkin spaces by the results of [Johnsen 1996]. Composition with the operator $Q_{1,+}$ preserves this ellipticity, so $\{P'_+, \gamma_0\}$ forms an elliptic system with regards to the mapping property

$$\{P'_+, \gamma_0\} : \begin{cases} \overline{F}_{p,q}^s(\Omega) \to \overline{F}_{p,q}^{s-1}(\Omega) \times B_{p,p}^{s-1/p}(\partial\Omega), \\ \overline{B}_{p,q}^s(\Omega) \to \overline{B}_{p,q}^{s-1}(\Omega) \times B_{p,q}^{s-1/p}(\partial\Omega), \end{cases} \tag{4-14}$$

for $s > 1/p$. Hence there is a parametrix

$$\begin{pmatrix} R'_D & K'_D \end{pmatrix}$$

of this Dirichlet problem, continuous in the opposite direction of (4-14). This shows (1).

Next, we can discuss the general problem (4-5) by the help of this special problem; such a discussion is standard within the Boutet de Monvel calculus. Define S'_B by (4-10), it is a ψdo on $\partial\Omega$ of order m_0 by the rules of calculus. If it is elliptic, it has a parametrix, which we denote $\widetilde{S}'_B$.

On the principal symbol level, the discussion takes place for exact operators; here we denote principal symbols of the involved operators P', B', K'_D, etc., by p', b', k'_D, etc. To solve the model problem (at a point (x', ξ') with $\xi' \neq 0$), with $g \in L_2(\mathbb{R}_+)$, $\psi \in \mathbb{C}$,

$$p'_+(x', \xi', D_n)w(x_n) = g(x_n) \text{ on } \mathbb{R}_+, \quad \gamma_0 b'_+(x', \xi', D_n)w(x_n) = \psi \text{ at } x_n = 0, \tag{4-15}$$

let $z = w - r'_D g$; then z should satisfy

$$p'_+ z = 0, \quad \gamma_0 b'_+ z = \psi - \gamma_0 b'_+ r'_D g \equiv \zeta. \tag{4-16}$$

Assuming that z satisfies the first equation, set

$$\gamma_0 z = \varphi; \text{ then } z = k'_D \varphi,$$

as the solution of the semihomogeneous Dirichlet problem for p'_+. To adapt z to the second part of (4.16), we require that $\gamma_0 b'_+ z = \zeta$; here

$$\gamma_0 b'_+ z = \gamma_0 b'_+ k'_D \varphi = s'_B \varphi,$$

when we define s'_B by (4-10) on the principal symbol level; it is just a complex number depending on (x', ξ'). The equation

$$s'_B \varphi = \zeta \tag{4-17}$$

is uniquely solvable precisely when $s'_B \neq 0$. In that case, (4-17) is solved uniquely by $\varphi = (s'_B)^{-1}\zeta$.

With this choice of φ, $z = k'_D \varphi$ is the unique solution of (4-16), and $w = r'_D g + z$ is the unique solution of (4-15). The formula in complete detail is

$$w = r'_D g + k'_D (s'_B)^{-1}\zeta = (I - k'_D (s'_B)^{-1}\gamma_0 b'_+)r'_D g + k'_D (s'_B)^{-1}\psi. \tag{4-18}$$

Expressed for the full operators, this shows that the problem (4-5) is elliptic precisely when the ψdo S'_B is so.

For the full operators, a similar construction can be carried out in a parametrix sense, but it is perhaps simpler to test directly by compositions that the operator

$$\begin{pmatrix} R'_B & K'_B \end{pmatrix} = \begin{pmatrix} (I - K'_D \widetilde{S}'_B \gamma_0 B'_+)R'_D & K'_D \widetilde{S}'_B \end{pmatrix}, \tag{4-19}$$

defined in analogy with (4-18), is a parametrix for $\{P'_+, \gamma_0 B'_+\}$: since $R'_D P'_+ + K'_D \gamma_0 = I + \mathcal{R}$ and $\widetilde{S}'_B \gamma_0 B'_+ K_D = \widetilde{S}'_B S'_B = I + \mathcal{S}$, with operators $\mathcal{R}$ and $\mathcal{S}$ of order $-\infty$, we have

$$
\begin{aligned}
\begin{pmatrix} R'_B & K'_B \end{pmatrix} \begin{pmatrix} P'_+ \\ \gamma_0 B'_+ \end{pmatrix} &= (I - K'_D \widetilde{S}'_B \gamma_0 B'_+) R'_D P'_+ + K'_D \widetilde{S}'_B \gamma_0 B'_+ \\
&= (I - K'_D \widetilde{S}'_B \gamma_0 B'_+)(1 + \mathcal{R} - K'_D \gamma_0) + K'_D \widetilde{S}'_B \gamma_0 B'_+ \\
&= I - K'_D \widetilde{S}'_B \gamma_0 B'_+ - K'_D \gamma_0 + K'_D \widetilde{S}'_B \gamma_0 B'_+ K'_D \gamma_0 + K'_D \widetilde{S}'_B \gamma_0 B'_+ + \mathcal{R}_1 \\
&= I + \mathcal{R}_2,
\end{aligned}
\tag{4-20}
$$

with operators $\mathcal{R}_1$ and $\mathcal{R}_2$ of order $-\infty$. The composition in the opposite order is similarly checked.

All this takes place in the Boutet de Monvel calculus. For our original problem we now find the parametrix as in (4-11), by the transition described in Theorem 4.1. $\qquad\square$

The order assumption on B was made for the sake of arriving at operators to which the Boutet de Monvel calculus applies. We think that m_0 could be allowed to be noninteger, with some more effort, drawing on results from [Grubb and Hörmander 1990].

The treatment can be extended to problems with vector-valued boundary conditions $\gamma_0 r^+ B$, where we also involve $\varrho_{\mu,M}$ for $M > 1$; see (3-8).

4B. *The Neumann boundary operator $\gamma_{\mu_0-1,1}$.* For ease of comparison to [Grubb 2015a], we denote the μ used above by μ_0 here.

The boundary conditions with B of noninteger order $m_0 + \mu_0$ are generally nonlocal, since B is so. But there do exist local boundary conditions too. For example, the Dirichlet-type operator $\gamma_{\mu_0-1,0}$ is local; see (2-23). So are the systems (see (3-8)) $\varrho_{\mu_0-M,M} = \{\gamma_{\mu_0-M,0}, \dots, \gamma_{\mu_0-M,M-1}\}$, which also define Fredholm operators together with $r^+ P$; see Theorem 3.2(3). Note that $\{r^+ P, \varrho_{\mu_0-M,M}\}$ operates on a larger space $X_{p,q}^{(\mu_0-M)(s)}(\overline{\Omega})$ than $X_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega})$ when $M > 1$.

What we shall show now is that one can impose a *higher-order local* boundary condition defined on $X_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega})$ *itself*, leading to a meaningful boundary value problem with Fredholm solvability under a reasonable ellipticity condition.

Here we treat the Neumann-type condition $\gamma_{\mu_0-1,1} u = \psi$, recalling from [Grubb 2015a, (5.3)ff.] that

$$
\gamma_{\mu_0-1,1} u = \Gamma(\mu_0 + 1) \gamma_0 \big(\partial_n (d(x)^{1-\mu_0} u) \big).
\tag{4-21}
$$

By application of (3-8) with $M = 2$, $\mu = \mu_0 - 1$,

$$
\gamma_{\mu_0-1,1} = \gamma_{\mu,M-1} : \begin{cases} F_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega}) \to B_{p,p}^{s-\operatorname{Re}\mu_0-1/p}(\partial\Omega), \\ B_{p,q}^{(\mu_0-1)(s)}(\overline{\Omega}) \to B_{p,q}^{s-\operatorname{Re}\mu_0-1/p}(\partial\Omega), \end{cases}
\tag{4-22}
$$

is well-defined for $s > \operatorname{Re}\mu + M - 1/p' = \operatorname{Re}\mu_0 + 1/p$.

The discussion of ellipticity takes place in local coordinates, so let us now assume that we are in a localized situation where P is given on $\mathbb{R}^n$, globally estimated, elliptic of order m and of type μ_0 and with factorization index μ_0 relative to the subset $\mathbb{R}^n_+$, as in [Grubb 2015a, Theorem 6.5].

For $\mathbb{R}^n_+$, we can express $\gamma_{\mu_0-1,1}$ in terms of auxiliary operators by

$$\gamma_{\mu_0-1,1}u = \gamma_0\partial_n \Xi_+^{\mu_0-1}u - (\mu_0-1)[D']\gamma_0 \Xi_+^{\mu_0-1}u; \tag{4-23}$$

see the calculations after Corollary 5.3 in [Grubb 2015a]. (In the manifold situation there is a certain freedom in choosing $d(x)$ and ∂_n, so we are tacitly assuming that a choice has been made that carries over to $d(x) = x_n$, $\partial_n = \partial/\partial x_n$ in the localization.)

There is an obstacle to applying the results of Section 4A to this, namely, that $\Xi_+^{\mu_0-1}$ is not truly a ψdo! This is a difficult fact that has been observed throughout the development of the theory. However, in connection with boundary conditions, operators like Ξ_+^μ work to some extent like the truly pseudodifferential operators Λ_+^μ. It is for this reason that we gave two versions of the operator K_D in (4-2) and the following, stemming from [Grubb 2015a, Theorem 6.5], in which Lemma 6.6 there was used.

Theorem 4.3. *Let P be given on $\mathbb{R}^n$, globally estimated, elliptic of order m and of type μ_0 and with factorization index μ_0 relative to the subset $\mathbb{R}^n_+$, and let $\begin{pmatrix} R_D & K_D \end{pmatrix}$ be a parametrix of the nonhomogeneous Dirichlet problem, as recalled in (4-2) and the following, with $K_D = \Xi_+^{1-\mu_0}e^+K'$ for a certain Poisson operator K' of order 0.*

Consider the Neumann-type problem

$$r^+Pu = f, \quad \gamma_{\mu_0-1,1}u = \psi, \tag{4-24}$$

where

$$\{r^+P, \gamma_{\mu_0-1,1}\} : \begin{cases} F_{p,q}^{(\mu_0-1)(s)}(\overline{\mathbb{R}}^n_+) \to \overline{F}_{p,q}^{s-\mathrm{Re}\,m}(\mathbb{R}^n_+) \times B_{p,p}^{s-\mathrm{Re}\,\mu_0-1/p}(\mathbb{R}^{n-1}), \\ B_{p,q}^{(\mu_0-1)(s)}(\overline{\mathbb{R}}^n_+) \to \overline{B}_{p,q}^{s-\mathrm{Re}\,m}(\mathbb{R}^n_+) \times B_{p,q}^{s-\mathrm{Re}\,\mu_0-1/p}(\mathbb{R}^{n-1}), \end{cases} \tag{4-25}$$

for $s > \mu_0 + 1/p$.

(1) *The operator*

$$S_N = \gamma_{\mu_0-1,1}K_D \tag{4-26}$$

equals $(\gamma_0\partial_n - (\mu_0-1)[D']\gamma_0)K'$ and is a ψdo on $\mathbb{R}^{n-1}$ of order 1.

(2) *If S_N is elliptic, then, with a parametrix of S_N denoted $\widetilde{S}_N$, there is the parametrix for $\{r^+P, \gamma_{\mu_0-1,1}\}$*

$$\begin{pmatrix} R_N & K_N \end{pmatrix} = \begin{pmatrix} (I - K_D\widetilde{S}_N\gamma_{\mu_0-1,1})R_D & K_D\widetilde{S}_N \end{pmatrix}. \tag{4-27}$$

(3) *Ellipticity holds in particular when the principal symbol of P equals $c(x)|\xi|^{2\mu_0}$, with $\mathrm{Re}\,\mu_0 > 0$, $c(x) \neq 0$.*

Proof. (1) By the formulas for $\gamma_{\mu_0-1,1}$ and K_D,

$$S_N = \gamma_{\mu_0-1,1}K_D = (\gamma_0\partial_n - (\mu_0-1)[D']\gamma_0)\Xi_+^{\mu_0-1}\Xi_+^{1-\mu_0}K' = (\gamma_0\partial_n - (\mu_0-1)[D']\gamma_0)K',$$

and it follows from the rules of the Boutet de Monvel calculus that this is a ψdo on $\mathbb{R}^{n-1}$ of order 1.

(2) In the elliptic case, one checks that (4-27) is a parametrix by calculations as in Theorem 4.2.

(3) In this case, the model problem for $\{r^+P, \gamma_{\mu_0-1,1}\}$ can be reduced to that for $\{r^+(1-\Delta)^{\mu_0}, \gamma_{\mu_0-1,1}\}$. For the latter, we have shown unique solvability in Theorem A.2 and Remark A.3 in the appendix. $\square$

Remark 4.4. The operator S_N is in fact the *Dirichlet-to-Neumann* operator for P, sending the Dirichlet data over into the Neumann data for solutions of $r^+ P u = 0$ in an approximate sense (modulo operators of order $-\infty$). From the calculations in the Appendix we see that its principal symbol equals $-\mu_0|\xi'|$ when P is principally equal to $(-\Delta)^{\mu_0}$, with $\mathrm{Re}\,\mu_0 > 0$.

4C. *Systems, further perspectives.* The factorization property used above will not in general hold for systems ($N \times N$-matrices) in a convenient way with smooth dependence on ξ', even if every element of the matrix has a factorization. But with the μ-transmission property we can establish an extremely useful connection to systems in the Boutet de Monvel calculus:

Proposition 4.5. *Let N be an integer ≥ 1, and let P be an $N \times N$-system, $P = (P_{jk})_{j,k=1,\ldots,N}$, of classical ψdo's P_{jk} of order $m \in \mathbb{C}$ on Ω_1 and of type $\mu \in \mathbb{C}$ relative to Ω. Let $\mu_0 \in \mu + \mathbb{Z}$. Then the operator*

$$Q = \Lambda_-^{(\mu_0 - m)} P \Lambda_+^{(-\mu_0)} \tag{4-28}$$

is of order and type 0, and hence belongs to the Boutet de Monvel calculus.

Proof. The factors $\Lambda_-^{(\mu_0 - m)}$ and $\Lambda_+^{(-\mu_0)}$ should be understood as diagonal matrices with $\Lambda_-^{(\mu_0 - m)}$ and $\Lambda_+^{(-\mu_0)}$, respectively, in the diagonal. When they are composed with P, they act on each entry by defining an operator of order and type 0 by the symbol composition rules. $\qquad\square$

This will allow for a general application of the Boutet de Monvel theory in the discussion of boundary value problems. Leaving the most general case for future works, we shall in the present paper just draw conclusions for systems where the operator (4-28) defines a system Q_+ that is in itself elliptic. Let us give a name to such cases, where the present considerations will apply without further efforts:

Definition 4.6. Let N be an integer ≥ 1, and let P be an elliptic $N \times N$-system, $P = (P_{jk})_{j,k=1,\ldots,N}$, of classical ψdo's P_{jk} of order $m \in \mathbb{C}$ on Ω_1 and of type $\mu \in \mathbb{C}$ relative to Ω. Let $\mu_0 \in \mu + \mathbb{Z}$. Then P is said to be μ_0-*reducible* when the operator Q, defined in (4-28) of order and type 0, has the property that Q_+ *is elliptic in the Boutet de Monvel calculus* (without auxiliary boundary operators).

The condition in the definition means that in local coordinates at the boundary, the model operator $q_0(x', 0, \xi', D_n)_+$ is bijective in $L_2(\mathbb{R}_+)^N$. It holds for $N = 1$ for the operators with factorization index μ_0, as accounted for in the proof of [Grubb 2015a, Theorem 4.4]. Another important case is where the operator P (a scalar or a system) is strongly elliptic, as observed in [Eskin 1981, Example 17.1].

Lemma 4.7. *Let $N \geq 1$, and let P be of order $m \in \mathbb{R}_+$ on Ω_1 and of type $\mu_0 = m/2$ relative to Ω. If P is strongly elliptic, i.e., satisfies in local coordinates (with $c > 0$),*

$$\mathrm{Re}(p_0(x, \xi)v, v) \geq c|\xi|^m |v|^2 \text{ for all } \xi \in \mathbb{R}^n, \, v \in \mathbb{C}^N,$$

then P is μ_0-reducible.

Proof. Here Q equals $\Lambda_-^{(-m/2)} P \Lambda_+^{(-m/2)}$. This is strongly elliptic of order 0, because the principal symbols of $\Lambda_-^{(-m/2)}$ and $\Lambda_+^{(-m/2)}$ are conjugates and homogeneous elliptic of order $-m/2$:

$$\mathrm{Re}(q_0(x, \xi)v, v) = \mathrm{Re}\big(p_0(x, \xi)\lambda_{+,0}^{-m/2}(\xi)v, \, \lambda_{+,0}^{-m/2}(\xi)v\big) \geq c|\xi|^m |\lambda_{+,0}^{-m/2}(\xi)v|^2 \geq c'|v|^2,$$

for all $\xi \in \mathbb{R}^n$, $v \in \mathbb{C}^N$, in local coordinates. Thus for each $x' \in \partial\Omega$, $\xi' \neq 0$, the model operator $q_0(x', 0, \xi', D_n)$ on $\mathbb{R}$ satisfies

$$\mathrm{Re}(q_0 u, u) \geq C \|u\|^2_{L_2(\mathbb{R})^N} \quad \text{for } u \in L_2(\mathbb{R})^N,$$

as seen by Fourier transformation in ξ_n. In particular, the restriction of $r^+ q_0$ to $C_0^\infty(\mathbb{R}_+)^N$ satisfies the above inequality, and the inequality extends to its closure, $r^+ q_0 e^+$, defined on $L_2(\mathbb{R}_+)^N$, which is therefore injective. Similar considerations hold for the adjoint, so indeed, $q_0(x', 0, \xi', D_n)_+$ is bijective in $L_2(\mathbb{R}_+)^N$. $\qquad\square$

Theorem 4.8. *Let P be an elliptic $N \times N$ system, $P = (P_{jk})_{j,k=1,\ldots,N}$, of classical ψdo's P_{jk} of order $m \in \mathbb{C}$ on Ω_1 and of type $\mu_0 \in \mathbb{C}$ relative to Ω.*

Define Q by (4-28) and assume that P is μ_0-reducible. Then we have:

(1) *Let $s > \mathrm{Re}\,\mu_0 - 1/p'$. If $u \in \dot{X}^\sigma_{p,q}(\overline{\Omega})^N$ for some $\sigma > \mathrm{Re}\,\mu_0 - 1/p'$ and $r^+ Pu \in \overline{X}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N$, then $u \in X^{\mu_0(s)}_{p,q}(\overline{\Omega})^N$. The mapping*

$$r^+ P : X^{\mu_0(s)}_{p,q}(\overline{\Omega})^N \to \overline{X}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N \tag{4-29}$$

is Fredholm, and has the parametrix

$$R = \Lambda_+^{(-\mu_0)} e^+ \widetilde{Q_+} \Lambda_{-,+}^{(\mu_0-m)} : \overline{X}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N \to X^{\mu_0(s)}_{p,q}(\overline{\Omega})^N, \tag{4-30}$$

where $\widetilde{Q_+}$ is a parametrix of Q_+. It has the structure $\widetilde{Q}_+ + G$ with G a singular Green operator of order and class 0.

(2) *In particular, if $r^+ Pu \in C^\infty(\overline{\Omega})^N$, then $u \in \mathscr{E}_{\mu_0}(\overline{\Omega})^N$, and the mapping*

$$r^+ P : \mathscr{E}_{\mu_0}(\overline{\Omega})^N \to C^\infty(\overline{\Omega})^N \tag{4-31}$$

is Fredholm.

(3) *Moreover, let $\mu = \mu_0 - M$ for a positive integer M. Then when $s > \mathrm{Re}\,\mu_0 - 1/p'$, $\{r^+ P, \varrho_{\mu,M}\}$ defines a Fredholm operator*

$$\{r^+ P, \varrho_{\mu,M}\} : \begin{cases} F^{\mu(s)}_{p,q}(\overline{\Omega})^N \to \overline{F}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N \times \prod_{0 \leq j < M} B^{s-\mathrm{Re}\,\mu-j-1/p}_{p,p}(\partial\Omega)^N, \\ B^{\mu(s)}_{p,q}(\overline{\Omega})^N \to \overline{B}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N \times \prod_{0 \leq j < M} B^{s-\mathrm{Re}\,\mu-j-1/p}_{p,q}(\partial\Omega)^N. \end{cases} \tag{4-32}$$

Proof. The proof goes as in [Grubb 2015a, Theorems 4.4 and 6.1]:

(1) We replace the equation

$$r^+ Pu = f \in \overline{X}^{s-\mathrm{Re}\,m}_{p,q}(\Omega)^N, \tag{4-33}$$

by composition on the left with $\Lambda_{-,+}^{(\mu_0-m)}$, by the equivalent problem

$$\Lambda_{-,+}^{(\mu_0-m)} r^+ Pu = g, \quad \text{where } g = \Lambda_{-,+}^{(\mu_0-m)} f \in \overline{X}^{s-\mathrm{Re}\,\mu_0}_{p,q}(\Omega)^N, \tag{4-34}$$

using the homeomorphism properties of $\Lambda_{-,+}^{(\mu_0-m)}$, applied to vectors. Here $f = \Lambda_{-,+}^{(m-\mu_0)}g$. Moreover (see Remark 1.1 in [Grubb 2015a]),

$$\Lambda_{-,+}^{(\mu_0-m)}r^+Pu = r^+\Lambda_-^{(\mu_0-m)}Pu.$$

Next, we set $v = r^+\Lambda_+^{(\mu_0)}u$; then $u = \Lambda_+^{(-\mu_0)}e^+v$, and equation (4-33) becomes

$$Q_+v = g, \quad \text{with } g \text{ given in } \overline{X}_{p,q}^{s-\mathrm{Re}\,\mu_0}(\Omega), \tag{4-35}$$

where Q is defined by (4-28).

The properties of P imply that Q is elliptic of order 0 and type 0, and hence belongs to the Boutet de Monvel calculus. The rest of the argument takes place within that calculus. By our assumption, $Q_+ = r^+Qe^+$ defines an elliptic boundary problem (without auxiliary trace or Poisson operators) there, and Q_+ is continuous in $\overline{X}_{p,q}^t(\Omega)$ for $t > -1/p'$. By the ellipticity, Q_+ has a parametrix $\widetilde{Q_+}$, continuous in the opposite direction, and with the mentioned structure. Since $v \in \dot{X}_{p,q}^{-1/p'+0}(\overline{\Omega})$ by hypothesis, solutions of $Q_+v = g$ with $g \in \overline{X}_{p,q}^t(\Omega)$ for some $t > -1/p'$ are in $\overline{X}_{p,q}^t(\Omega)$. Moreover,

$$Q_+ : \overline{X}_{p,q}^t(\Omega) \to \overline{X}_{p,q}^t(\Omega) \quad \text{is Fredholm for all } t > -1/p'.$$

When carried back to the original functions, this shows (1).

(2) This follows by letting $s \to \infty$, using that $\bigcap_s X_{p,q}^{\mu(s)}(\overline{\Omega})^N = \mathscr{E}_\mu(\overline{\Omega})^N$.

(3) We use that the mapping $\varrho_{\mu,M}$ in (3-8) extends immediately to vector-valued functions

$$\varrho_{\mu,M} : \begin{cases} F_{p,q}^{\mu(s)}(\overline{\Omega})^N \to \prod_{0\le j<M} B_{p,p}^{s-\mathrm{Re}\,\mu-j-1/p}(\partial\Omega)^N, \\ B_{p,q}^{\mu(s)}(\overline{\Omega})^N \to \prod_{0\le j<M} B_{p,p}^{s-\mathrm{Re}\,\mu-j-1/p}(\partial\Omega)^N, \end{cases} \tag{4-36}$$

when $s > \mathrm{Re}\,\mu_0 - 1/p'$, surjective with null-space $X_{p,q}^{\mu_0(s)}(\overline{\Omega})^N$ (recall $\mu = \mu_0 - M$). When we adjoin this mapping to (4-29), we obtain (4-32).   □

One of the things we obtain here is that results from [Eskin 1981] (extended to L_p in [Shargorodsky 1994; Chkadua and Duduchava 2001]), on solvability for s in an interval of length 1 around $\mathrm{Re}\,\mu_0$, are lifted to regularity and Fredholm properties for all larger s, with exact information on the domain, also in general scales of function spaces. Moreover, our theorem is obtained via a systematic variable-coefficient calculus, whereas the results in [Eskin 1981] are derived from constant-coefficient considerations by ad hoc perturbation methods in L_2-Sobolev spaces.

Also the results on other boundary conditions in the present paper extend to suitable systems. One can moreover extend the results to operators in vector bundles (since they can be locally expressed by matrices of operators).

The Boutet de Monvel theory is not an easy theory (as the elaborate presentations [Boutet de Monvel 1971; Rempel and Schulze 1982; Grubb 1984; 1990; 1996; 2009; Schrohe 2001] in the literature show), but one could have feared that a theory for the more general μ-transmission operators and their boundary problems would be a step up in difficulty. Fortunately, as we have seen, many of the issues can be dealt with by reductions using the special operators $\Lambda_\pm^{(\mu)}$ to cases where the type 0 theory applies.

There is currently also an interest in problems with less smooth symbols. For this connection, we mention that there do exist pseudodifferential theories for such problems, also with boundary conditions; see [Abels 2005; Grubb 2014] and their references. One finds that a lack of smoothness in the x-variable narrows down the interval of parameters s where one has good solvability properties, and compositions are delicate. It is also possible to work under limitations on the number of standard estimates in ξ.

Appendix: Calculations in an explicit example

Pseudodifferential methods are a refinement of the application of the Fourier transform, making it useful even for variable-coefficient partial differential operators, and, for example, allowing generalizations to operators of noninteger order. But to explain some basic mechanisms, it may be useful to consider a simple "constant-coefficient" case, where explicit elementary calculations can be made, not requiring intricate composition rules. This is the case for $(1 - \Delta)^a$ ($a > 0$) on $\mathbb{R}^n_+$, where everything can be worked out by hand in exact detail (in the spirit of the elementary [Grubb 2009, Chapter 9]). We here restrict the attention to H^s_p-spaces.

The symbol of $(1 - \Delta)^a$ factors as

$$(\langle \xi' \rangle^2 + \xi_n^2)^a = (\langle \xi' \rangle - i\xi_n)^a (\langle \xi' \rangle + i\xi_n)^a. \tag{A-1}$$

Now we shall use the definitions of simple order-reducing operators $\Xi^t_\pm$ and Poisson operators K_j from [Grubb 2015a], with $\langle \xi' \rangle$ instead of $[\xi']$, because they fit particularly well with the factors in (A-1). We shall often abbreviate $\langle \xi' \rangle$ to σ.

The *homogeneous Dirichlet problem*

$$r^+ (1 - \Delta)^a u = f, \quad \text{with } f \text{ given in } \overline{H}^{s-2a}_p(\mathbb{R}^n_+), \tag{A-2}$$

$s > a - 1/p'$, has a unique solution u in $\dot{H}^{a-1/p'+0}_p(\overline{\mathbb{R}}^n_+)$ determined as follows:

With $\Xi^t_\pm = \text{OP}((\langle \xi' \rangle + i\xi_n)^t)$, we have that $(1 - \Delta)^a = \Xi^a_- \Xi^a_+$ on $\mathbb{R}^n$. Let $v = r^+ \Xi^a_+ u$; it is in $\overline{H}^{-1/p'+0}_p(\mathbb{R}^n_+) = \dot{H}^{-1/p'+0}_p(\overline{\mathbb{R}}^n_+)$, and $u = \Xi^{-a}_+ e^+ v$. Then (A-2) becomes

$$r^+ \Xi^a_- e^+ v = f. \tag{A-3}$$

Here $r^+ \Xi^a_- e^+ = \Xi^a_{-,+}$ is known to map $\overline{H}^t_p(\mathbb{R}^n_+)$ homeomorphically onto $\overline{H}^{t-a}_p(\mathbb{R}^n_+)$ for all $t \in \mathbb{R}$, with inverse $\Xi^{-a}_{-,+}$ (see, e.g., [Grubb 2015a, Section 1].) In particular, with f given in $\overline{H}^{s-2a}_p(\mathbb{R}^n_+)$, (A-3) has the unique solution $v = \Xi^{-a}_{-,+} f \in \overline{H}^{s-a}_p(\mathbb{R}^n_+)$. Then (A-2) has the unique solution

$$u = \Xi^{-a}_+ e^+ \Xi^{-a}_{-,+} f \equiv R_D f, \tag{A-4}$$

and it belongs to $H^{a(s)}_p(\overline{\mathbb{R}}^n_+)$ by the definition of that space. Thus the solution operator for (A-2) is $R_D = \Xi^{-a}_+ e^+ \Xi^{-a}_{-,+}$. (This is a simple variant of the proof of [Grubb 2015a, Theorem 4.4].)

Next, we go to the larger space $H^{(a-1)(s)}_p(\overline{\mathbb{R}}^n_+)$, still assuming $s > a - 1/p'$, where we study the *nonhomogeneous Dirichlet problem*. By [Grubb 2015a, Theorem 5.1] with $\mu = a - 1$ and $M = 1$, we

have a mapping $\gamma_{a-1,0}$, acting as

$$\gamma_{a-1,0} : u \mapsto \Gamma(a)\gamma_0(x_n^{1-a}u),$$

also equal to $\gamma_0 \Xi_+^{a-1} u$, and sending $H_p^{(a-1)(s)}(\bar{\mathbb{R}}_+^n)$ onto $B_p^{s-a+1-1/p}(\mathbb{R}^{n-1})$ with kernel $H_p^{a(s)}(\bar{\mathbb{R}}_+^n)$. Together with $(1-\Delta)^a$, it therefore defines a homeomorphism for $s > a - 1/p'$,

$$\{r^+(1-\Delta)^a, \gamma_{a-1,0}\} : H_p^{(a-1)(s)}(\bar{\mathbb{R}}_+^n) \to \bar{H}_p^{s-2a}(\bar{\mathbb{R}}_+^n) \times B_p^{s-a+1-1/p}(\mathbb{R}^{n-1}). \tag{A-5}$$

It represents the problem

$$r^+(1-\Delta)^a u = f, \quad \gamma_{a-1,0}u = \varphi, \tag{A-6}$$

which we regard as the nonhomogeneous Dirichlet problem for $(1-\Delta)^a$. The solution operator in the case $\varphi = 0$ is clearly R_D defined above, since the kernel of $\gamma_{a-1,0}$ is $H_p^{a(s)}(\bar{\mathbb{R}}_+^n)$.

Also, the solution operator for the problem (A-6) with $f = 0$ can be found explicitly:

On the boundary symbol level we consider the problem (recall $\sigma = \langle \xi' \rangle$)

$$(\sigma - \partial_n)^a (\sigma + \partial_n)^a u(x_n) = 0 \quad \text{on } \mathbb{R}_+. \tag{A-7}$$

Since $\text{OP}_n((\sigma - i\xi_n)^\mu)$ preserves support in $\bar{\mathbb{R}}_-$ for all μ, u must equivalently satisfy

$$(\sigma + \partial_n)^a u(x_n) = 0 \quad \text{on } \mathbb{R}_+. \tag{A-8}$$

This has the distribution solution

$$u(x_n) = \mathscr{F}_{\xi_n \to x_n}^{-1}(\sigma + i\xi_n)^{-a} = \Gamma(a)^{-1} x_n^{a-1} e^+ r^+ e^{-\sigma x_n} \tag{A-9}$$

(see, e.g., [Hörmander 1983, Example 7.1.17] or [Grubb 2015a, (2.5)]), and the derivatives $\partial_n^k u$ are likewise solutions, since

$$(\sigma + i\xi_n)^a (i\xi_n)^k (\sigma + i\xi_n)^{-a} = (i\xi_n)^k = \mathscr{F}_{x_n \to \xi_n} \delta_0^{(k)},$$

where $\delta_0^{(k)}$ is supported in $\{0\}$. The undifferentiated function matches our problem. Set

$$\tilde{k}_{a-1,0}(x_n, \xi') = \Gamma(a)^{-1} x_n^{a-1} e^+ r^+ e^{-\sigma x_n} = \mathscr{F}_{\xi_n \to x_n}^{-1}(\sigma + i\xi_n)^{-a}; \tag{A-10}$$

then, since $\gamma_{a-1,0}\tilde{k}_{a-1,0} = 1$, the mapping $\mathbb{C} \ni \varphi \mapsto \varphi \cdot r^+ \tilde{k}_{a-1,0}$ solves the problem

$$(\sigma + \partial_n)^a u(x_n) = 0 \text{ on } \mathbb{R}_+, \quad \gamma_{a-1,0}u = \varphi. \tag{A-11}$$

Using the Fourier transform in ξ' also, we find that (A-6) with $f = 0$ has the solution

$$u(x) = K_{a-1,0}\varphi \equiv \mathscr{F}_{\xi' \to x'}^{-1}(\tilde{k}_{a-1,0}(x_n, \xi')\hat{\varphi}(\xi')). \tag{A-12}$$

It can be denoted $\text{OPK}(\tilde{k}_{a-1,0})\varphi$, by a generalization of the notation from the Boutet de Monvel calculus. We moreover define $k_{a-1,0}(\xi) = \mathscr{F}_{x_n \to \xi_n}\tilde{k}_{a-1,0}(x_n, \xi') = (\sigma + i\xi_n)^{-a}$; $\tilde{k}_{a-1,0}$ and $k_{a-1,0}$ are the symbol-kernel and symbol of $K_{a-1,0}$, respectively.

Note that

$$k_{a-1,0}(\xi', \xi_n) = (\langle\xi'\rangle + i\xi_n)^{-a} = (\langle\xi'\rangle + i\xi_n)^{1-a}(\langle\xi'\rangle + i\xi_n)^{-1}, \quad \text{hence}$$

$$K_{a-1,0} = \Xi_+^{1-a} K_0, \tag{A-13}$$

where $K_0 = \text{OPK}\big((\langle\xi'\rangle + i\xi_n)^{-1}\big)$ is *the Poisson operator for the Dirichlet problem for* $1 - \Delta$,

$$K_0\varphi = \mathscr{F}_{\xi\to x}^{-1}((\langle\xi'\rangle + i\xi_n)^{-1}\hat\varphi(\xi'))$$

(see, e.g., [Grubb 2009, Chapter 9]). K_0 is usually considered as mapping into a space over $\mathbb{R}_+^n$, and it is well-known that $K_0 : B_p^{t-1/p}(\mathbb{R}^{n-1}) \to \overline{H}_p^t(\mathbb{R}_+^n)$ for all $t \in \mathbb{R}$. However, the above formula shows that it in fact maps into distributions on $\mathbb{R}^n$ supported in $\overline{\mathbb{R}}_+^n$, so we can, with a slight abuse of notation, identify K_0 with e^+K_0, mapping into $e^+\overline{H}_p^t(\mathbb{R}_+^n)$, and conclude that

$$K_{a-1,0} : B_p^{s-a+1-1/p}(\mathbb{R}^{n-1}) \to H_p^{(a-1)(s)}(\overline{\mathbb{R}}_+^n) \quad \text{for all } s \in \mathbb{R}. \tag{A-14}$$

We have shown:

Theorem A.1. *Let* $a > 0$. *The nonhomogeneous Dirichlet problem* (A-6) *for* $(1-\Delta)^a$ *on* $\mathbb{R}_+^n$ *is uniquely solvable, in that the operator* (A-5) *for* $s > a - 1/p'$ *has inverse*

$$\begin{pmatrix} r^+(1-\Delta)^a \\ \gamma_{a-1,0} \end{pmatrix}^{-1} = \begin{pmatrix} R_D & K_{a-1,0} \end{pmatrix}, \tag{A-15}$$

where R_D *and* $K_{a-1,0}$ *are defined in* (A-4) *and* (A-12).

Third, we consider the boundary problem

$$r^+(1-\Delta)^a u = f, \quad \gamma_{a-1,1}u = \psi, \tag{A-16}$$

which we shall view as a *nonhomogeneous Neumann problem* for $(1 - \Delta)^a$. We here assume $s > (a-1) + 2 - 1/p' = a + 1/p$, to use the construction in [Grubb 2015a, Theorem 5.1] with $\mu = a - 1$, $M = 2$. Recall from [Grubb 2015a, (5.3)ff.], that $\gamma_{a-1,1}$ acts as

$$\gamma_{a-1,1} : u \mapsto \Gamma(a+1)\gamma_0(\partial_n(x_n^{1-a}u)). \tag{A-17}$$

Moreover, we can infer from the text after Corollary 5.3 in [Grubb 2015a] (with $[\xi']$ replaced by $\langle\xi'\rangle$) that

$$\gamma_{a-1,1}u = \gamma_0\partial_n\Xi_+^{a-1}u - (a-1)\langle D'\rangle\gamma_{a-1,0}u$$

for $u \in H_p^{(a-1)(s)}(\overline{\mathbb{R}}_+^n)$ with $s > a + 1/p$. Then, for a null solution z written in the form $z = K_{a-1,0}\varphi = \Xi_+^{1-a}K_0\varphi$ (recall (A-13)), we have, since $\gamma_0\partial_n K_0 = -\langle D'\rangle$,

$$\gamma_{a-1,1}z = \gamma_0\partial_n\Xi_+^{a-1}z - (a-1)\langle D'\rangle\gamma_{a-1,0}z = \gamma_0\partial_n K_0\varphi - (a-1)\langle D'\rangle\varphi = -a\langle D'\rangle\varphi.$$

Hence in order for z to solve (A-16) with $f = 0$, φ must satisfy

$$\psi = -a\langle D'\rangle\varphi.$$

Since $a \neq 0$, the coefficient $-a\langle D' \rangle$ is an elliptic invertible ψdo, so (A-16) with $f = 0$ is uniquely solvable with solution

$$z = K_N \psi, \quad \text{where } K_N = -K_{a-1,0} a^{-1} \langle D' \rangle^{-1} = -\Xi_+^{1-a} K_0 a^{-1} \langle D' \rangle^{-1}. \tag{A-18}$$

To solve (A-16) with a given $f \neq 0$, and $\psi = 0$, we let $v = R_D f$ and reduce to the problem for $z = u - v$:

$$r^+ (1 - \Delta)^a (u - v) = 0, \quad \gamma_{a-1,1}(u - v) = -\gamma_{a-1,1} R_D f.$$

This has the unique solution

$$u - v = -K_N \gamma_{a-1,1} R_D f; \quad \text{and hence } u = R_D f - K_N \gamma_{a-1,1} R_D f.$$

Altogether, we find:

Theorem A.2. *The Neumann problem* (A-16) *for* $(1 - \Delta)^a$ *on* $\mathbb{R}_+^n$ *is uniquely solvable, in that the operator*

$$\{r^+(1 - \Delta)^a, \gamma_{a-1,1}\} : H_p^{(a-1)(s)}(\overline{\mathbb{R}}_+^n) \to \overline{H}_p^{s-2a}(\mathbb{R}_+^n) \times B_p^{s-a-1/p}(\mathbb{R}^{n-1}), \tag{A-19}$$

for $s > a + 1/p$ *is a homeomorphism, with inverse*

$$\begin{pmatrix} R_N & K_N \end{pmatrix} = \begin{pmatrix} (I - K_N \gamma_{a-1,1}) R_D & K_N \end{pmatrix}, \tag{A-20}$$

with R_D *and* K_N *described in* (A-4) *and* (A-18).

Note that there is here a *Dirichlet-to-Neumann operator* P_{DN} sending the Dirichlet-type data over into Neumann-type data for solutions of $r^+(1 - \Delta)^a u = 0$:

$$P_{DN} = -a\langle D' \rangle. \tag{A-21}$$

Remark A.3. We have here assumed a real in order to relate to the fractional powers of the Laplacian, but all the above goes through in the same way if a is replaced by a complex μ with $\operatorname{Re} \mu > 0$; then in Sobolev exponents and inequalities for s, a should be replaced by $\operatorname{Re} \mu$.

One can also let higher order boundary operators $\gamma_{a-1,j}$ enter in a similar way, defining single boundary conditions.

Acknowledgement

References

[Abatangelo 2013] N. Abatangelo, "Large *s*-harmonic functions and boundary blow-up solutions for the fractional Laplacian", preprint, 2013. arXiv 1310.3193

[Abels 2005] H. Abels, "Pseudodifferential boundary value problems with non-smooth coefficients", *Comm. Partial Differential Equations* **30**:10-12 (2005), 1463–1503. MR 2006m:35395 Zbl 1087.35099

[Albin and Melrose 2009] P. Albin and R. Melrose, "Fredholm realizations of elliptic symbols on manifolds with boundary", *J. Reine Angew. Math.* **627** (2009), 155–181. MR 2011e:58046 Zbl 1161.58011

[Caffarelli and Silvestre 2009] L. Caffarelli and L. Silvestre, "Regularity theory for fully nonlinear integro-differential equations", *Comm. Pure Appl. Math.* **62**:5 (2009), 597–638. MR 2010d:35376 Zbl 1170.45006

[Chkadua and Duduchava 2001] O. Chkadua and R. Duduchava, "Pseudodifferential equations on manifolds with boundary: Fredholm property and asymptotic", *Math. Nachr.* **222** (2001), 79–139. MR 2001m:58057 Zbl 0985.47019

[Eskin 1981] G. I. Eskin, *Boundary value problems for elliptic pseudodifferential equations*, Translations of Mathematical Monographs **52**, American Mathematical Society, Providence, R.I., 1981. Translated from the Russian by S. Smith. MR 82k:35105 Zbl 0458.35002

[Felsinger et al. 2014] M. Felsinger, M. Kassmann, and P. Voigt, "The Dirichlet problem for nonlocal operators", *Math. Zeitschrift* (online publication November 2014).

[Frank and Geisinger 2014] R. L. Frank and L. Geisinger, "Refined semiclassical asymptotics for fractional powers of the Laplace operator", *J. Reine Angew. Math.* (online publication January 2014).

[Grubb 1984] G. Grubb, "Singular Green operators and their spectral asymptotics", *Duke Math. J.* **51**:3 (1984), 477–528. MR 86d:35109 Zbl 0553.58034

[Grubb 1990] G. Grubb, "Pseudo-differential boundary problems in L_p spaces", *Comm. Partial Differential Equations* **15**:3 (1990), 289–340. MR 91c:47108 Zbl 0723.35091

[Grubb 1996] G. Grubb, *Functional calculus of pseudodifferential boundary problems*, 2nd ed., Progress in Mathematics **65**, Birkhäuser, Boston, MA, 1996. MR 96m:35001 Zbl 0844.35002

[Grubb 2009] G. Grubb, *Distributions and operators*, Graduate Texts in Mathematics **252**, Springer, New York, 2009. MR 2010b:46081 Zbl 1171.47001

[Grubb 2014] G. Grubb, "Spectral asymptotics for nonsmooth singular Green operators", *Comm. Partial Differential Equations* **39**:3 (2014), 530–573. With an appendix by H. Abels. MR 3169794 Zbl 1290.35057

[Grubb 2015a] G. Grubb, "Fractional Laplacians on domains, a development of Hörmander's theory of μ-transmission pseudodifferential operators", *Adv. Math.* **268** (2015), 478–528. MR 3276603 Zbl 06370011

[Grubb 2015b] G. Grubb, "Spectral results for mixed problems and fractional elliptic operators", *J. Math. Anal. Appl.* **421**:2 (2015), 1616–1634. MR 3258341 Zbl 06347113

[Grubb and Hörmander 1990] G. Grubb and L. Hörmander, "The transmission property", *Math. Scand.* **67**:2 (1990), 273–289. MR 92f:58175 Zbl 0766.35088

[Harutyunyan and Schulze 2008] G. Harutyunyan and B.-W. Schulze, *Elliptic mixed, transmission and singular crack problems*, EMS Tracts in Mathematics **4**, European Mathematical Society (EMS), Zürich, 2008. MR 2009k:35002 Zbl 1141.35001

[Hoh and Jacob 1996] W. Hoh and N. Jacob, "On the Dirichlet problem for pseudodifferential operators generating Feller semigroups", *J. Funct. Anal.* **137**:1 (1996), 19–48. MR 97c:35227 Zbl 0901.47027

[Hörmander 1965] L. Hörmander, "Chapter II: Boundary problems for "classical" pseudo-differential operators", unpublished lecture notes at Inst. Adv. Study, Princeton, 1965, Available at http://www.math.ku.dk/~grubb/LH65.pdf.

[Hörmander 1983] L. Hörmander, *The analysis of linear partial differential operators, I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **256**, Springer, Berlin, 1983. MR 85g:35002a Zbl 0521.35001

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **274**, Springer, Berlin, 1985. MR 87d:35002a Zbl 0601.35001

[Hörmander 1997] L. Hörmander, *Lectures on nonlinear hyperbolic differential equations*, Mathématiques & Applications **26**, Springer, Berlin, 1997. MR 98e:35103 Zbl 0881.35001

[Johnsen 1996] J. Johnsen, "Elliptic boundary problems and the Boutet de Monvel calculus in Besov and Triebel–Lizorkin spaces", *Math. Scand.* **79**:1 (1996), 25–85. MR 97m:35294 Zbl 0873.35023

[Melrose 1993] R. B. Melrose, *The Atiyah–Patodi–Singer index theorem*, Research Notes in Mathematics **4**, A K Peters, Ltd., Wellesley, MA, 1993. MR 96g:58180 Zbl 0796.58050

[Boutet de Monvel 1971] L. Boutet de Monvel, "Boundary problems for pseudo-differential operators", *Acta Math.* **126**:1-2 (1971), 11–51. MR 53 #11674 Zbl 0206.39401

[Rempel and Schulze 1982] S. Rempel and B.-W. Schulze, *Index theory of elliptic boundary problems*, Akademie-Verlag, Berlin, 1982. MR 85b:58126 Zbl 0504.35002

[Rempel and Schulze 1984] S. Rempel and B.-W. Schulze, "Complex powers for pseudodifferential boundary problems, II", *Math. Nachr.* **116** (1984), 269–314. MR 86b:35207b Zbl 0585.58041

[Ros-Oton and Serra 2014a] X. Ros-Oton and J. Serra, "Boundary regularity for fully nonlinear integro-differential equations", preprint, 2014, Available at http://arxiv.org/abs/1404.1197. arXiv 1404.1197

[Ros-Oton and Serra 2014b] X. Ros-Oton and J. Serra, "The Dirichlet problem for the fractional Laplacian: regularity up to the boundary", *J. Math. Pures Appl.* (9) **101**:3 (2014), 275–302. MR 3168912 Zbl 1285.35020

[Schrohe 2001] E. Schrohe, "A short introduction to Boutet de Monvel's calculus", pp. 85–116 in *Approaches to Singular Analysis*, edited by J. Gil et al., Operator Theory: Advances and Applications **125**, Birkhäuser Basel, 2001. Zbl 0985.35122

[Shargorodsky 1994] E. Shargorodsky, "An L_p-analogue of the Vishik–Eskin theory", *Memoirs on Differential Equations and Mathematical Physics* **2** (1994), 41–146.

[Taylor 1981] M. E. Taylor, *Pseudodifferential operators*, Princeton Mathematical Series **34**, Princeton University Press, 1981. MR 82i:35172 Zbl 0453.47026

[Triebel 1995] H. Triebel, *Interpolation theory, function spaces, differential operators*, 2nd ed., Johann Ambrosius Barth,, Heidelberg, 1995. MR 96f:46001 Zbl 0830.46028

GERD GRUBB: grubb@math.ku.dk
Department of Mathematical Sciences, Copenhagen University, Universitetsparken 5, DK-2100 Copenhagen, Denmark

ON THE UNCONDITIONAL UNIQUENESS OF SOLUTIONS
TO THE INFINITE RADIAL CHERN–SIMONS–SCHRÖDINGER HIERARCHY

Xuwen Chen and Paul Smith

In this article, we establish the unconditional uniqueness of solutions to an infinite radial Chern–Simons–Schrödinger (IRCSS) hierarchy in two spatial dimensions. The IRCSS hierarchy is a system of infinitely many coupled PDEs that describes the limiting Chern–Simons–Schrödinger dynamics of infinitely many interacting anyons. The anyons are two-dimensional objects that interact through a self-generated field. Due to the interactions with the self-generated field, the IRCSS hierarchy is a system of *nonlinear* PDEs, which distinguishes it from the *linear* infinite hierarchies studied previously. Factorized solutions of the IRCSS hierarchy are determined by solutions of the Chern–Simons–Schrödinger system. Our result therefore implies the unconditional uniqueness of solutions to the radial Chern–Simons–Schrödinger system as well.

1. Introduction

1A. *The Chern–Simons–Schrödinger system.* The Chern–Simons–Schrödinger system is given by

$$\begin{cases} D_t\phi = i\sum_{\ell=1}^2 D_\ell D_\ell \phi + ig|\phi|^2\phi, \\ \partial_t A_1 - \partial_1 A_0 = -\operatorname{Im}(\bar\phi D_2\phi), \\ \partial_t A_2 - \partial_2 A_0 = \operatorname{Im}(\bar\phi D_1\phi), \\ \partial_1 A_2 - \partial_2 A_1 = -\tfrac{1}{2}|\phi|^2, \end{cases} \tag{1}$$

where the associated covariant differentiation operators are defined in terms of the potential A by

$$D_\alpha := \partial_\alpha + iA_\alpha, \quad \alpha \in \{0,1,2\}, \tag{2}$$

and where we adopt the convention that $\partial_0 := \partial_t$ and $D_t := D_0$. The wavefunction ϕ is complex-valued, the potential A a real-valued 1-form, and the pair (A, ϕ) is defined on $I \times \mathbb{R}^2$ for some time interval I. The Lagrangian action for this system is

$$L(A, \phi) = \frac{1}{2} \int_{\mathbb{R}^{2+1}} \left[\text{Im}(\bar{\phi} D_t \phi) + |D_x \phi|^2 - \frac{g}{2} |\phi|^4 \right] dx\, dt + \frac{1}{2} \int_{\mathbb{R}^{2+1}} A \wedge dA, \tag{3}$$

where here $|D_x \phi|^2 := |D_1 \phi|^2 + |D_2 \phi|^2$. Although the potential A appears explicitly in the Lagrangian, it is easy to see that locally $L(A, \phi)$ only depends upon the field $F = dA$. Precisely, the Lagrangian is invariant with respect to the gauge transformations

$$\phi \mapsto e^{-i\theta} \phi, \qquad A \mapsto A + d\theta \tag{4}$$

for compactly supported real-valued functions $\theta(t, x)$. The Chern–Simons–Schrödinger system (1), obtained as the Euler–Lagrange equations of (3), inherits this gauge freedom.

The system (1) is a basic model of Chern–Simons dynamics [Jackiw and Pi 1992; Ezawa, Hotta, and Iwazaki 1991a; 1991b; Jackiw and Pi 1991]. It plays a role in describing certain physical phenomena, such as the fractional quantum Hall effect, high-temperature superconductivity, and Aharonov–Bohm scattering, and also provides an example of a Galilean-invariant planar gauge field theory [Jackiw and Templeton 1981; Deser, Jackiw, and Templeton 1982; Jackiw, Pi, and Weinberg 1991; Martina, Pashaev, and Soliani 1993; Wilczek 1990].

One interpretation of (1) is as a mean-field equation. Informally, one may consider (1) as describing the behavior of a large number of anyons, interacting with each other directly and through a self-generated field, in the case where the N-body wave function factorizes. There are a number of challenges one encounters in trying to formalize and prove this statement, and this paper addresses some of them. We will postpone further discussion of many-body dynamics to the next subsection and instead point out that, because the main evolution equation in (1) includes a cubic nonlinearity, one might hope to prove for (1) what one can prove for the cubic nonlinear Schrödinger equation (NLS). It is important to note, however, that (1) has many nonlinear terms, some nonlocal and some involving the derivative of the wave function. These terms appear because of the geometric structure that arises from modeling the interactions with the self-generated field. Due to the complexity of the nonlinearity in (1) and the gauge freedom (4), the system (1) is significantly more challenging to analyze than the cubic NLS. This difference is seen even at the level of the wellposedness theory, to which we now turn.

The system (1) is Galilean-invariant and has conserved *charge*

$$\text{chg}(\phi) := \int_{\mathbb{R}^2} |\phi|^2 \, dx \tag{5}$$

and *energy*

$$E(\phi) := \frac{1}{2} \int_{\mathbb{R}^2} \left[|D_x \phi|^2 - \frac{g}{2} |\phi|^4 \right] dx. \tag{6}$$

Moreover, for each $\lambda > 0$, there is the scaling symmetry

$$\phi(t, x) \mapsto \lambda \phi(\lambda^2 t, \lambda x), \qquad A_j(t, x) \mapsto \lambda A_j(\lambda^2 t, \lambda x), \quad j \in \{1, 2\},$$

$$\phi_0(x) \mapsto \lambda \phi_0(\lambda x), \qquad A_0(t, x) \mapsto \lambda^2 A_0(\lambda^2 t, \lambda x),$$

which preserves both the system and the charge of the initial data ϕ_0. Therefore, from the point of view of wellposedness theory, the system (1) is L^2-critical. We remark that system (1) is defocusing when $g < 1$ and focusing when $g \geq 1$. The defocusing/focusing dichotomy is most readily seen by rewriting the energy (6) using the so-called Bogomol'nyi identity. After using this identity, one may also see the dichotomy manifested in the virial and Morawetz identities. For more details, see [Liu and Smith 2014, §4, §5]. Note also that the sign convention for g that we adopt, which is the one used in the Chern–Simons literature, is opposite to the usual one adopted for the cubic NLS. A more significant difference between Chern–Simons systems and the cubic NLS is that, unlike the case for the cubic NLS, the coupling parameter g cannot be rescaled to belong to a discrete set of canonical values.

Nevertheless, (1) is ill-posed so long as it retains the gauge freedom (4). This freedom is eliminated by imposing an additional constraint equation. The most common gauge choice for studying (1) is the Coulomb gauge, which is the constraint

$$\partial_1 A_1 + \partial_2 A_2 = 0. \tag{7}$$

Coupling (7) with the field equations quickly leads to explicit expressions for A_α, $\alpha = 0, 1, 2$, in terms of ϕ. These expressions also happen to be nonlinear and nonlocal:

$$A_0 = \Delta^{-1}[\partial_1 \operatorname{Im}(\bar{\phi} D_2 \phi) - \partial_2 \operatorname{Im}(\bar{\phi} D_1 \phi)], \qquad A_1 = \tfrac{1}{2}\Delta^{-1}\partial_2|\phi|^2, \qquad A_2 = -\tfrac{1}{2}\Delta^{-1}\partial_1|\phi|^2. \tag{8}$$

Local wellposedness of (1) with respect to the Coulomb gauge at the Sobolev regularity of H^2 is established in [Bergé, De Bouard, and Saut 1995]. This is improved to H^1 in [Huh 2013]. Local wellposedness for data small in H^s, $s > 0$, is established in [Liu, Smith, and Tataru 2012] using the heat gauge, whose defining condition is $\partial_1 A_1 + \partial_2 A_2 = A_0$. This result relies upon various Strichartz-type spaces as well as more sophisticated U^p and V^p spaces. We refer the reader to [Liu, Smith, and Tataru 2012, §2] for a comparison of the Coulomb and heat gauges.

In symmetry-reduced settings, one may say more, and in particular, [Liu and Smith 2014] establishes large-data global wellposedness results at the critical regularity for the equivariant Chern–Simons–Schrödinger system. To introduce the equivariance (or *vortex*) ansatz, it is convenient to use polar coordinates. Define

$$A_r = \frac{x_1}{|x|} A_1 + \frac{x_2}{|x|} A_2, \qquad A_\theta = -x_2 A_1 + x_1 A_2. \tag{9}$$

We can invert the transform by writing

$$A_1 = A_r \cos\theta - \tfrac{1}{r} A_\theta \sin\theta, \qquad A_2 = A_r \sin\theta + \tfrac{1}{r} A_\theta \cos\theta. \tag{10}$$

Note that these relations are analogous to

$$\partial_r = \frac{x_1}{|x|} \partial_1 + \frac{x_2}{|x|} \partial_2, \qquad \partial_\theta = -x_2 \partial_1 + x_1 \partial_2$$

and

$$\partial_1 = (\cos\theta)\partial_r - \tfrac{1}{r}(\sin\theta)\partial_\theta, \qquad \partial_2 = (\sin\theta)\partial_r + \tfrac{1}{r}(\cos\theta)\partial_\theta.$$

The equivariant ansatz, then, is

$$\phi(t, x) = e^{im\theta}u(t, r), \quad A_1(t, x) = -\frac{x_2}{r}v(t, r), \quad A_2(t, x) = \frac{x_1}{r}v(t, r), \quad A_0(t, x) = w(t, r), \quad (11)$$

where we assume that m is a nonnegative integer, u is real-valued at time zero, and v and w are real-valued for all time. This ansatz implies that $A_r = 0$ and that A_θ is a radial function. It also places us in the Coulomb gauge, i.e., $\partial_1 A_1 + \partial_2 A_2 = 0$ or equivalently $\partial_r A_r + \frac{1}{r} A_r + \frac{1}{r^2} \partial_\theta A_\theta = 0$. For some motivation for studying vortex solutions in Chern–Simons theories, see [Paul and Khare 1986; de Vega and Schaposnik 1986a; 1986b; Jackiw and Weinberg 1990; R. M. Chen and Spirn 2009; Byeon, Huh, and Seok 2012].

Converting (1) into polar coordinates and utilizing (11), we obtain the equivariant Chern–Simons–Schrödinger system (see [Liu and Smith 2014, §1] for full details):

$$\begin{cases} (i\partial_t + \Delta)\phi = \frac{2m}{r^2} A_\theta \phi + A_0 \phi + \frac{1}{r^2} A_\theta^2 \phi - g|\phi|^2\phi, \\ \partial_r A_0 = \frac{1}{r}(m + A_\theta)|\phi|^2, \\ \partial_t A_\theta = r \, \text{Im}(\bar{\phi}\partial_r \phi), \\ \partial_r A_\theta = -\frac{1}{2}|\phi|^2 r, \\ A_r = 0. \end{cases} \quad (12)$$

Global wellposedness holds for equivariant L^2 data of arbitrary (nonnegative) charge in the defocusing case $g < 1$ and for L^2 data with charge less than that of the ground state in the focusing case $g \geq 1$; this is the main result of [Liu and Smith 2014].

In this paper, we are interested in the radial case ($m = 0$) of system (12), which is

$$\begin{cases} (i\partial_t + \Delta)\phi = A_0 \phi + \frac{1}{r^2} A_\theta^2 \phi - g|\phi|^2\phi, \\ \partial_r A_0 = \frac{1}{r} A_\theta |\phi|^2, \\ \partial_t A_\theta = r \, \text{Im}(\bar{\phi}\partial_r \phi), \\ \partial_r A_\theta = -\frac{1}{2}|\phi|^2 r, \\ A_r = 0. \end{cases} \quad (13)$$

1B. *The infinite Chern–Simons–Schrödinger hierarchy.* The infinite Chern–Simons–Schrödinger hierarchy is a sequence of trace class nonnegative operator kernels that are symmetric in the sense that

$$\gamma^{(k)}(t, \boldsymbol{x}_k, \boldsymbol{x}_k') = \overline{\gamma^{(k)}(t, \boldsymbol{x}_k', \boldsymbol{x}_k)},$$

and

$$\gamma^{(k)}(t, x_{\sigma(1)}, \ldots, x_{\sigma(k)}, x_{\sigma(1)}', \ldots, x_{\sigma(k)}') = \gamma^{(k)}(t, x_1, \ldots, x_k, x_1', \ldots, x_k'), \quad (14)$$

for any permutation σ, and which satisfy the two-dimensional infinite Chern–Simons–Schrödinger hierarchy of equations

$$\partial_t \gamma^{(k)} + \sum_{j=1}^{k} [iA_0(t, x_j), \gamma^{(k)}] = \sum_{j=1}^{k} \sum_{\ell=1}^{2} i[D_{x_j^{(\ell)}} D_{x_j^{(\ell)}}, \gamma^{(k)}] + ig \sum_{j=1}^{k} B_{j,k+1} \gamma^{(k+1)}, \quad (15)$$

where $\mathbb{R}^2 \ni x_j = (x_j^{(1)}, x_j^{(2)})$ for each j, as well as the corresponding field-current identities from [Jackiw and Pi 1990, (1.7a)–(1.7c)], i.e.,

$$
\begin{cases}
F_{01} = -P_2(t, x) - A_2(t, x)\rho(t, x), \\
F_{02} = P_1(t, x) + A_1(t, x)\rho(t, x), \\
F_{12} = -\frac{1}{2}\rho(t, x),
\end{cases}
\tag{16}
$$

where, as before, $F := dA$. Here g is the coupling constant,

$$
B_{j,k+1}\gamma^{(k+1)} := \mathrm{Tr}_{k+1}[\delta(x_j - x_{k+1}), \gamma^{(k+1)}],
\tag{17}
$$

the momentum $P(t, x)$ is given by

$$
P(t, x) := \int e^{i(\xi-\xi')x}\frac{\xi+\xi'}{2}\hat{\gamma}^{(1)}(t, \xi, \xi')\, d\xi\, d\xi',
$$

and $\rho(t, x)$ is a shorthand for

$$
\rho(t, x) := \gamma^{(1)}(t, x, x).
\tag{18}
$$

Each $x_j \in \mathbb{R}^2$, and $\boldsymbol{x}_k := (x_1, \ldots, x_k) \in \mathbb{R}^{2k}$. Given a compactly supported $\theta(t, x)$, the kernels $\gamma^{(k)}$ and potential A transform under a change of gauge according to

$$
\gamma^{(k)} \mapsto \gamma^{(k)} \prod_{j=1}^{k} e^{-i\theta(t,x_j)}e^{i\theta(t,x_j')}, \qquad A \mapsto A + d\theta.
$$

The invariance of (15) and (16) under such transformations can be checked straightforwardly.

For the purposes of our analysis, it is more convenient to write (15) as

$$
i\partial_t \gamma^{(k)} + \sum_{j=1}^{k}[\Delta_{x_j}, \gamma^{(k)}] = \sum_{j=1}^{k}\sum_{\ell=1}^{2}[-2iA_{x_j^{(\ell)}}\partial_{x_j^{(\ell)}} - i\partial_{x_j^{(\ell)}}A_{x_j^{(\ell)}} + A_{x_j^{(\ell)}}^2, \gamma^{(k)}]
$$
$$
+ \sum_{j=1}^{k}[A_0(t, x_j), \gamma^{(k)}] - g\sum_{j=1}^{k}B_{j,k+1}\gamma^{(k+1)}. \tag{19}
$$

The Coulomb gauge condition (7), upon being coupled to (16), leads to

$$
A_0 = \Delta^{-1}[\partial_1(P_2 + A_2\rho) - \partial_2(P_1 + A_1\rho)], \qquad A_1 = \tfrac{1}{2}\Delta^{-1}\partial_2\rho, \qquad A_2 = -\tfrac{1}{2}\Delta^{-1}\partial_1\rho.
$$

This is analogous to how (8) for the Chern–Simons–Schrödinger system (1) is obtained by coupling to the field equations in (1) the gauge condition (7). Because each A_α involves ρ, defined in (18), it is clear that each term involving $\gamma^{(k)}$ in the right-hand side of (19) is best thought of as a nonlinear term. This nonlinear dependence persists under changes of gauge, though some gauges lead to tamer nonlinearities than others.

We remark that, while the specific form the nonlinearity of (19) takes indeed depends upon the gauge selection made, the observables associated with the system do not depend upon the gauge choice.

We note that the system (1) generates a special solution to the infinite hierarchy (15)–(16). In particular,

if (A, ϕ) solves (1), then $(A, \{\gamma^{(k)}\})$ solves (15)–(16), where each $\gamma^{(k)}$ is given by

$$\gamma^{(k)}(t, \boldsymbol{x}_k, \boldsymbol{x}'_k) = \prod_{j=1}^{k} \phi(t, x_j) \bar{\phi}(t, x'_j).$$

We start our analysis of many-body dynamics with the above infinite hierarchy. Ideally, one would prefer instead to begin with a many-body system with only finitely many quantum particles. Because the basic particles in question are neither bosons nor fermions, there are difficulties to overcome with such an approach. Concerning the difficulties in dealing with microscopic statistics, one can refer to [Benedetto, Castella, Esposito, and Pulvirenti 2005], for instance. Fortunately, as remarked in [Benedetto, Castella, Esposito, and Pulvirenti 2005], microscopic statistics disappear as the particle number tends to infinity. Thus, the infinite hierarchy satisfies the symmetry condition (14). We finally remark that the field equations (16) depend merely on the 1-particle density $\gamma^{(1)}$, as has been observed formally in the physics literature [Deser, Jackiw, and Templeton 1982; Jackiw, Pi, and Weinberg 1991; Jackiw and Pi 1991; Jackiw and Templeton 1981; Jackiw and Weinberg 1990].

One motivation for pursuing an analysis of the infinite hierarchy even without first specifying the finite hierarchy is that the known approaches to rigorously deriving mean-field equations, e.g., the Boltzmann equation and the cubic NLS, all require a uniqueness theorem for the corresponding infinite hierarchy. Establishing uniqueness of the infinite hierarchy is, moreover, a critical step. We therefore anticipate that our result in this article will be the linchpin of any future rigorous derivation of the Chern–Simons–Schrödinger system.

As remarked before, the analysis of the Chern–Simons–Schrödinger system with general data is, at the moment, very delicate. The same remark applies all the more to the associated infinite hierarchy, to which (1) is a special solution. Thus, we consider the radial version of the infinite Chern–Simons–Schrödinger hierarchy in this paper. The nonradial equivariant case ($m > 0$), though still much simpler than the general system, is slightly more challenging than the radial case. Unfortunately, the techniques we employ for studying the radial case do not immediately extend to the nonradial equivariant case due to certain logarithmic divergences.

The infinite radial Chern–Simons–Schrödinger hierarchy. The Chern–Simons–Schrödinger system (1) simplifies to (13) under the assumption of radiality. Similarly, by assuming radiality, we reduce Equations (15) through (18) to the infinite radial Chern–Simons–Schrödinger hierarchy

$$i\partial_t \gamma^{(k)} + \sum_{j=1}^{k} [\Delta_{x_j}, \gamma^{(k)}] = \sum_{j=1}^{k} \Big[A_0(t, |x_j|) + \frac{1}{|x_j|^2} A_\theta^2(t, |x_j|), \gamma^{(k)} \Big] - g \sum_{j=1}^{k} B_{j,k+1} \gamma^{(k+1)} \qquad (20)$$

and the field equations

$$F_{r\theta}(t, |x|) = -\tfrac{1}{2}|x|\rho(t, |x|)$$

and

$$F_{0\theta}(t, |x|) = |x| P_r(t, |x|),$$
$$F_{0r}(t, |x|) = -\tfrac{1}{|x|} A_\theta(t, |x|) \rho(t, |x|),$$

for $\gamma^{(k)} = \gamma^{(k)}(t, \boldsymbol{r}_k, \boldsymbol{r}'_k)$. In particular, here we assume that

$$
\begin{aligned}
\gamma^{(k)} &= u(t, \boldsymbol{r}_k, \boldsymbol{r}'_k), \\
A_r &= 0, \\
A_\theta &= v(t, r),
\end{aligned}
$$

where u is real-valued at time zero and v is real-valued for all time. This assumption enforces the Coulomb gauge. Recall that $B_{j,k+1}$ is defined in (17) and ρ is given by (18). As before, $F := dA$, though now we are adopting polar coordinates for A. Though we could rewrite everything exclusively in terms of polar coordinates, we choose instead to use both Cartesian and polar coordinates.

Putting everything together, we see that we are studying solutions $\gamma^{(k)} = \gamma^{(k)}(t, \boldsymbol{r}_k, \boldsymbol{r}'_k)$ of

$$
\begin{cases}
i\partial_t \gamma^{(k)} + \sum_{j=1}^{k} [\Delta_{x_j}, \gamma^{(k)}] = \sum_{j=1}^{k} \left[A_0(t, |x_j|) + \frac{1}{|x_j|^2} A_\theta^2(t, |x_j|), \gamma^{(k)} \right] - g \sum_{j=1}^{k} B_{j,k+1} \gamma^{(k+1)}, \\
\partial_r A_0(t, |x|) = \frac{1}{|x|} A_\theta \rho(t, |x|), \\
\partial_t A_\theta(t, |x|) = |x| P_r(t, |x|), \\
\partial_r A_\theta(t, |x|) = -\frac{1}{2} |x| \rho(t, |x|), \\
A_r = 0.
\end{cases}
\tag{21}
$$

We interpret $\gamma^{(k)}$ as a complex-valued function on $\mathbb{R}_t \times \mathbb{R}_+^k \times \mathbb{R}_+^k$ subject to the symmetries

$$
\gamma^{(k)}(t, \boldsymbol{r}_k, \boldsymbol{r}'_k) = \overline{\gamma^{(k)}(t, \boldsymbol{r}'_k, \boldsymbol{r}_k)}
$$

and

$$
\gamma^{(k)}(t, r_{\sigma(1)}, \ldots, r_{\sigma(k)}, r'_{\sigma(1)}, \ldots, r'_{\sigma(k)}) = \gamma^{(k)}(t, r_1, \ldots, r_k, r'_1, \ldots, r'_k).
\tag{22}
$$

Though each $r_j \in \mathbb{R}_+$, we associate to this space the measure $r\,dr$, as indeed we think of $r_j = |x_j|$ for $x_j \in \mathbb{R}^2$.

Note that we can eliminate A_θ and A_0 in (21). In particular, we have

$$
A_\theta(t, r) = -\frac{1}{2} \int_0^r \rho(t, s) s\, ds
\tag{23}
$$

and

$$
A_0(t, r) = \frac{1}{2} \int_r^\infty \rho(t, s) \int_0^s \rho(t, u) u\, du \, \frac{ds}{s},
\tag{24}
$$

which reflect the natural boundary conditions for A_θ and A_0 that we adopt for (1). Therefore, we may rewrite (21) as

$$
i\partial_t \gamma^{(k)} + \sum_{j=1}^{k} [\Delta_{x_j}, \gamma^{(k)}] = \sum_{j=1}^{k} \left[\frac{1}{2} \int_{r_j}^\infty \rho(t, s) \int_0^s \rho(t, u) u\, du \, \frac{ds}{s} + \frac{1}{r_j^2} \left(-\frac{1}{2} \int_0^{r_j} \rho(t, s) s\, ds \right)^2, \gamma^{(k)} \right]
$$

$$
- g \sum_{j=1}^{k} B_{j,k+1} \gamma^{(k+1)},
$$

$$
\gamma^{(k)}(0) = \gamma_0^{(k)}, \quad k \in \mathbb{N}.
\tag{25}
$$

1C. *Main results.* Our main theorem says that any admissible mild solution of the radial infinite CSS hierarchy is unconditionally unique in $L^\infty_{t\in[0,T)}\mathfrak{H}^{2/3}_{\mathrm{rad}}$. To explain what this means, for $s \in \mathbb{R}$, we define the space $\mathfrak{H}^s_{\mathrm{rad}}$ to be the collection of sequences $\{\gamma^{(k)}\}_{k\in\mathbb{N}}$ of density matrices in $L^2_{\mathrm{sym}}(\mathbb{R}^{2k})$ such that $\gamma^{(k)} = \gamma^{(k)}(t, \boldsymbol{r}_k, \boldsymbol{r}'_k)$ and

$$\mathrm{Tr}(|S^{(k,s)}\gamma^{(k)}|) < M^{2k} \quad \text{for all } k \in \mathbb{N} \text{ and for some constant } M > 0,$$

where

$$S^{(k,s)} := \prod_{j=1}^{k}(1 - \Delta_{x_j})^{s/2}(1 - \Delta_{x'_j})^{s/2}.$$

Here L^2_{sym} denotes the space of L^2 functions satisfying (14). Let $U^{(k)}(t)$ denote the propagator

$$U^{(k)}(t) := e^{it\Delta_{x_k}}e^{-it\Delta_{x'_k}}. \tag{26}$$

A *mild solution* of (25) in the space $L^\infty_{[0,T]}\mathfrak{H}^s_{\mathrm{rad}}$ is a sequence of marginal density matrices $\Gamma = (\gamma^{(k)}(t))_{k\in\mathbb{N}}$ solving

$$\gamma^{(k)}(t) = U^{(k)}(t)\gamma^{(k)}(0) - i\int_0^t U^{(k)}(t-s)\left(\sum_{j=1}^{k}\left[\frac{1}{2}\int_{r_j}^{\infty}\rho(t,v)\int_0^v \rho(t,u)u\,du\,\frac{dv}{v}\right.\right.$$
$$\left.\left. + \frac{1}{r_j^2}\left(-\frac{1}{2}\int_0^{r_j}\rho(t,v)v\,dv\right)^2, \gamma^{(k)}\right] - g\sum_{j=1}^{k}B_{j,k+1}\gamma^{(k+1)}\right)ds$$

and satisfying

$$\sup_{t\in[0,T)}\mathrm{Tr}(|S^{(k,s)}\gamma^{(k)}(t)|) < M^{2k}$$

for a finite constant M independent of k. Note that, if we are given factorized initial data

$$\gamma_0^{(k)}(\boldsymbol{r}_k, \boldsymbol{r}'_k) = \prod_{j=1}^{k}\phi_0(r_j)\overline{\phi_0(r'_j)},$$

then the condition that $(\gamma^{(k)}(0)) \in \mathfrak{H}^s_{\mathrm{rad}}$ is equivalent to

$$\mathrm{Tr}(|S^{(k,s)}\gamma^{(k)}(0)|) = \|\phi_0\|^{2k}_{H^s} < M^{2k}, \quad k \in \mathbb{N},$$

which is to say that $\|\phi_0\|_{H^s} < M$ for some $M < \infty$. Then a solution to the IRCSS hierarchy in $L^\infty_{t\in[0,T)}\mathfrak{H}^s_{\mathrm{rad}}$ is given by the sequence of factorized density matrices

$$\gamma^{(k)}(t, \boldsymbol{r}_k, \boldsymbol{r}'_k) = \prod_{j=1}^{k}\phi_t(r_j)\overline{\phi_t(r'_j)}$$

provided the corresponding 1-particle wave function ϕ_t satisfies the radial Chern–Simons–Schrödinger system (13).

Admissibility we take to mean that $\operatorname{Tr} \gamma^{(k)} = 1$ for all $k \in \mathbb{N}$ and

$$\gamma^{(k)} = \operatorname{Tr}_{k+1}(\gamma^{(k+1)}), \quad k \in \mathbb{N}. \tag{27}$$

This is required in our application of the quantum de Finetti theorem. As there are weak analogues of the quantum de Finetti theorem applicable to limiting hierarchies, we expect our techniques to apply to the problem of rigorously deriving the radial CSS from large, finite systems.

Theorem 1.1 (unconditional uniqueness for the infinite hierarchy). *There is at most one $L^\infty_{t \in [0,T)} \mathfrak{H}^{2/3}_{\mathrm{rad}}$ admissible solution to the infinite radial Chern–Simons–Schrödinger hierarchy* (21).

Theorem 1.2 (unconditional uniqueness for the Chern–Simons–Schrödinger system). *There is at most one $L^\infty_{t \in [0,T)} H^{2/3}(\mathbb{R}^2)$ solution to the radial Chern–Simons–Schrödinger system* (13).

Before explaining our main theorem, we first remark that deriving mean-field equations from many-body systems by studying infinite hierarchies is a very rich subject. For works related to the Boltzmann equation, see [Lanford 1975; King 1975; Arkeryd, Caprino, and Ianiro 1991; Cercignani, Illner, and Pulvirenti 1994; Gallagher, Saint-Raymond, and Texier 2013]. For works related to the Hartree equation, see [Spohn 1980; Fröhlich, Knowles, and Schwarz 2009; Erdős and Yau 2001; Rodnianski and Schlein 2009; Knowles and Pickl 2010; Grillakis, Machedon, and Margetis 2010; 2011; X. Chen 2012b; L. Chen, Lee, and Schlein 2011; Michelangeli and Schlein 2012; Ammari and Nier 2008; 2011; Lewin, Nam, and Rougerie 2014]. For works related to the cubic NLS, see [Adami, Golse, and Teta 2007; Elgart, Erdős, Schlein, and Yau 2006; Erdős, Schlein, and Yau 2006; 2007; 2010; 2009; Klainerman and Machedon 2008; Kirkpatrick, Schlein, and Staffilani 2011; T. Chen and Pavlović 2011; 2010; T. Chen, Pavlović, and Tzirakis 2012; T. Chen and Pavlović 2014; Pickl 2011; X. Chen 2012a; 2013; Benedikter, Oliveira, and Schlein 2012; Grillakis and Machedon 2013; X. Chen and Holmer 2013c; 2013b; T. Chen, Hainzl, Pavlović, and Seiringer 2014; X. Chen and Holmer 2013a; Hong, Taliaferro, and Xie 2014; Gressman, Sohinger, and Staffilani 2014; Sohinger and Staffilani 2014; Sohinger 2014a; 2014b]. For works related to the quantum Boltzmann equation, see [Benedetto, Castella, Esposito, and Pulvirenti 2006; 2005; 2008; 2004]. The infinite hierarchies considered previously to the present one are all linear. In contrast to this, the infinite radial Chern–Simons–Schrödinger hierarchy is nonlinear.

For our problem, we have taken the phrase "unconditional uniqueness" from the study of the NLS. It is shown by Cercignani's counterexample [Cercignani, Illner, and Pulvirenti 1994] that solutions to infinite hierarchies like the Boltzmann hierarchy and the Gross–Pitaevskii hierarchy are generally not unconditionally unique in the sense that a solution is not uniquely determined by the initial datum unless one assumes appropriate space-time bounds on the solution. In the NLS literature, "unconditional uniqueness" usually means establishing uniqueness without assuming that some Strichartz norm is finite. Since we are using tools from the study of the NLS, we therefore call our main theorems unconditional uniqueness theorems.[1]

---

[1]In other words, the uniqueness theorems regarding the Gross–Pitaevskii hierarchies [Klainerman and Machedon 2008; Kirkpatrick, Schlein, and Staffilani 2011; X. Chen 2012a; X. Chen and Holmer 2013a; Gressman, Sohinger, and Staffilani 2014] are conditional, whereas [Adami, Golse, and Teta 2007; Erdős, Schlein, and Yau 2007; T. Chen, Hainzl, Pavlović, and Seiringer 2014;

Finally, we remark that, for the proof of the main theorems, we apply the quantum de Finetti theorem in a manner similar to [T. Chen, Hainzl, Pavlović, and Seiringer 2014; Hong, Taliaferro, and Xie 2014] but with adjustments tailored to deal with the nonlinearity in the infinite hierarchy that we consider. The quantum de Finetti theorem is a version of the classical Hewitt–Savage theorem. T. Chen, C. Hainzl, N. Pavlović, and R. Seiringer are the first to apply the quantum de Finetti theorem to the study of infinite hierarchies in the quantum setting. For results regarding the uniqueness of the Boltzmann hierarchy using the Hewitt–Savage theorem, see [Arkeryd, Caprino, and Ianiro 1991].

2. Proof of the main theorem

We will prove that, if we are given two $L^\infty_{[0,T]} \mathfrak{H}^{2/3}_{\mathrm{rad}}$ solutions $\{\gamma_1^{(k)}\}$ and $\{\gamma_2^{(k)}\}$ to system (21) subject to the same initial datum, then the trace norm of the difference $\{\gamma^{(k)} = \gamma_1^{(k)} - \gamma_2^{(k)}\}$ is zero. In contrast to the usual infinite hierarchies (e.g., Boltzmann, Gross–Pitaevskii, ...), system (21) is nonlinear. Thus, $\gamma^{(k)}$ does not solve system (21). In order to show that $\gamma^{(k)}$ has zero trace norm, we first express $\gamma^{(k)}$ as a suitable Duhamel–Born series, which contains a nonlinear part and an interaction part (see Section 2A). These two parts we estimate separately with bounds contained respectively in Theorems 2.3 and 2.4, which together constitute our main estimates. In Section 2B, we prove the main theorem, Theorem 1.1, assuming the main estimates. The proof of Theorem 2.3 is postponed to Section 4 (and Theorem 2.4 we handle in this section).

2A. *Setup.* Set for short

$$a(r_j) := A_0(t, r_j) + \frac{1}{r_j^2} A_\theta^2(t, r_j) \tag{28}$$

and

$$a(\boldsymbol{r}_k) := \sum_{j=1}^{k} a(r_j). \tag{29}$$

Let $\mathscr{A}^{(k)}$ denote the operator that acts according to

$$\mathscr{A}^{(k)} f := [a(\boldsymbol{r}_k), f]. \tag{30}$$

Also, set for short

$$B_{k+1} := \sum_{j=1}^{k} B_{j,k+1} = \sum_{j=1}^{k} \mathrm{Tr}_{k+1}[\delta(x_j - x_{k+1}), \cdot]. \tag{31}$$

With these abbreviations, the first equation of (21) assumes the form

$$i\partial_t \gamma^{(k)} + [\Delta_{\boldsymbol{x}_k}, \gamma^{(k)}] = \mathscr{A}^{(k)} \gamma^{(k)} - g B_{k+1} \gamma^{(k+1)}. \tag{32}$$

Remark 2.1. The operator $\mathscr{A}^{(k)}$ is linear but itself depends upon $\gamma^{(1)}$. In fact, it only depends upon the diagonal $\rho(t, r) = \gamma^{(1)}(t, r, r)$. The term $\mathscr{A}^{(k)} \gamma^{(k)}$ is therefore better thought of as a nonlinear term rather than a linear one.

Hong, Taliaferro, and Xie 2014; Sohinger 2014b] are unconditional in the NLS sense. Yet they are all considered conditional in the Boltzmann literature.

Let $\{\gamma_1^{(k)}\}$ and $\{\gamma_2^{(k)}\}$ be solutions subject to the same initial data with, respectively, $\rho_1(t,r) := \gamma_1^{(1)}(t,r,r)$ and $\rho_2(t,r) := \gamma_2^{(1)}(t,r,r)$. Let $\gamma^{(k)} := \gamma_1^{(k)} - \gamma_2^{(k)}$. Then

$$i\partial_t \gamma^{(k)} + [\Delta_{\mathbf{x}_k}, \gamma^{(k)}] = \mathscr{A}_1^{(k)}\gamma_1^{(k)} - \mathscr{A}_2^{(k)}\gamma_2^{(k)} - gB_{k+1}\gamma^{(k+1)}. \tag{33}$$

We can rewrite (33) using the relation

$$\mathscr{A}_1^{(k)}\gamma_1^{(k)} - \mathscr{A}_2^{(k)}\gamma_2^{(k)} = \mathscr{A}_1^{(k)}\gamma^{(k)} + \mathscr{A}^{(k)}\gamma_2^{(k)},$$

where now

$$\mathscr{A}^{(k)} := \mathscr{A}_1^{(k)} - \mathscr{A}_2^{(k)},$$

so that it becomes

$$i\partial_t \gamma^{(k)} + [\Delta_{\mathbf{x}_k}, \gamma^{(k)}] = \mathscr{A}_1^{(k)}\gamma^{(k)} + \mathscr{A}^{(k)}\gamma_2^{(k)} - gB_{k+1}\gamma^{(k+1)} \tag{34}$$

or, equivalently,

$$(i\partial_t + \Delta_{\mathbf{x}_k} - \Delta_{\mathbf{x}_k'})\gamma^{(k)} = \mathscr{A}_1^{(k)}\gamma^{(k)} + \mathscr{A}^{(k)}\gamma_2^{(k)} - gB_{k+1}\gamma^{(k+1)}.$$

Recalling the corresponding linear propagator $U^{(k)}(t)$ defined in (26), we write (34) in integral form, i.e.,

$$\gamma^{(k)}(t_k) = -ig\int_0^{t_k} dt_{k+1} U^{(k)}(t_k - t_{k+1})\big[\mathscr{A}_1^{(k)}\gamma^{(k)}(t_{k+1}) + \mathscr{A}^{(k)}\gamma_2^{(k)}(t_{k+1}) + B_{k+1}\gamma^{(k+1)}(t_{k+1})\big]. \tag{35}$$

In invoking this formula in future calculations, we set $g = -1$ for simplicity and we ignore the i in front so that we do not need to keep track of its exact power, as the precise power is not relevant to the estimates.

Remark 2.2. The choice of $g = -1$ corresponds to a defocusing case in (12). It is important to note, however, that the choice $g = -1$ at this step is purely for the sake of convenience; all subsequent arguments can accommodate any $g \neq -1$ at the cost of certain powers of $|g|$. In particular, our arguments apply to the self-dual case $g = 1$, which is the most interesting from the physical point of view.

For the purpose of proving unconditional uniqueness, it suffices to show $\gamma^{(1)} = 0$. Iterating (35) l_c times,[2] we obtain

$$\gamma^{(1)}(t_1) = \int_0^{t_1} dt_2 U^{(1)}(t_1 - t_2)\big(\mathscr{A}_1^{(1)}\gamma^{(1)}(t_2) + \mathscr{A}^{(1)}\gamma_2^{(1)}(t_2)\big) + \int_0^{t_1} dt_2 U^{(1)}(t_1 - t_2)B_2\gamma^{(2)}(t_2)$$

$$= \int_0^{t_1} dt_2 U^{(1)}(t_1 - t_2)\big(\mathscr{A}_1^{(1)}\gamma^{(1)}(t_2) + \mathscr{A}^{(1)}\gamma_2^{(1)}(t_2)\big)$$

$$\qquad + \int_0^{t_1} dt_2 U^{(1)}(t_1 - t_2)B_2 \int_0^{t_2} dt_3 U^{(2)}(t_2 - t_3)\big(\mathscr{A}_1^{(2)}\gamma^{(2)}(t_3) + \mathscr{A}^{(2)}\gamma_2^{(2)}(t_3)\big)$$

$$\qquad + \int_0^{t_1} dt_2 U^{(1)}(t_1 - t_2)B_2 \int_0^{t_2} dt_3 U^{(2)}(t_2 - t_3)B_3\gamma^{(3)}(t_3)$$

$$= \cdots$$

$$= \mathrm{NP}^{(l_c)} + \mathrm{IP}^{(l_c)}, \tag{36}$$

---

[2]Here, l_c stands for the level of coupling. When $l_c = 0$, one recovers (34).

where $\mathrm{NP}^{(l_c)}$ and $\mathrm{IP}^{(l_c)}$, the nonlinear part and the interaction part, respectively, are given by

$$\mathrm{NP}^{(l_c)} = G^{(1)} + \sum_{r=1}^{l_c} \int_0^{t_1} \cdots \int_0^{t_r} dt_2 \cdots dt_{r+1} U^{(1)}(t_1 - t_2) B_2 \cdots U^{(r)}(t_r - t_{r+1}) B_{r+1} G^{(r+1)}(t_{r+1}) \quad (37)$$

and

$$\mathrm{IP}^{(l_c)} = \int_0^{t_1} \cdots \int_0^{t_{l_c+1}} dt_2 \cdots dt_{l_c+1} U^{(1)}(t_1 - t_2) B_2 \cdots U^{(l_c+1)}(t_{l_c} - t_{l_c+1}) B_{l_c+2} \gamma^{(l_c+2)}(t_{l_c+2}), \quad (38)$$

where

$$G^{(k)}(t_k) := \int_0^{t_k} dt_{k+1} U^{(k)}(t_k - t_{k+1}) \big(\mathcal{A}_1^{(k)} \gamma^{(k)}(t_{k+1}) + \mathcal{A}^{(k)} \gamma_2^{(k)}(t_{k+1}) \big).$$

2B. *Proof assuming the main estimates.*

Theorem 2.3. *There exists a constant $C > 0$ such that*

$$\mathrm{Tr}|\mathrm{NP}^{(l_c)}(t_1)| \le C t_1 \sup_{t \in [0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)|$$

for all coupling levels l_c and all sufficiently small t_1.

Proof. We postpone the proof to Section 3. □

Theorem 2.4. *There exists a constant $C > 0$ such that*

$$\mathrm{Tr}|\mathrm{IP}^{(l_c)}(t_1)| \le (C t_1^{1/3})^{l_c}$$

for all coupling levels l_c.

Proof. This estimate follows from the same method used for the corresponding term in [T. Chen, Hainzl, Pavlović, and Seiringer 2014], which relies on the quantum de Finetti theorem and on a combinatorial analysis of the graphs that one can associate to the Duhamel expansions. One merely needs to replace the three-dimensional trilinear estimates [T. Chen, Hainzl, Pavlović, and Seiringer 2014, (6.19), (6.20)] with (55) and (56), respectively, taking $s = \frac{2}{3}$, and replace the three-dimensional Sobolev estimate

$$\|f\|_{L^6(\mathbb{R}^3)} \lesssim \|f\|_{H^1(\mathbb{R}^3)}$$

with the two-dimensional Sobolev estimate

$$\|f\|_{L^6(\mathbb{R}^2)} \lesssim \|f\|_{H^{2/3}(\mathbb{R}^2)}.$$

We remark that it is because of this Sobolev estimate that we take $s = \frac{2}{3}$ in H^s rather than a smaller s. □

With Theorems 2.3 and 2.4, we then infer from (36) that

$$\mathrm{Tr}|\gamma^{(1)}(t_1)| \le \mathrm{Tr}|\mathrm{NP}^{(l_c)}(t_1)| + \mathrm{Tr}|\mathrm{IP}^{(l_c)}(t_1)|$$
$$\le C t_1 \sup_{t \in [0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| + (C t_1^{1/3})^{l_c}$$
$$\le C T \sup_{t \in [0,T]} \mathrm{Tr}|\gamma^{(1)}(t)| + (C T^{1/3})^{l_c}$$

for all $t_1 \in [0, T]$. Take the supremum in time on both sides to get

$$\sup_{t \in [0,T]} \mathrm{Tr} |\gamma^{(1)}(t)| \leq CT \sup_{t \in [0,T]} \mathrm{Tr} |\gamma^{(1)}(t)| + (CT^{1/3})^{l_c}.$$

Therefore, for all T small enough, we obtain

$$\tfrac{1}{2} \sup_{t \in [0,T]} \mathrm{Tr} |\gamma^{(1)}(t)| \leq (CT^{1/3})^{l_c} \to 0 \quad \text{as } l_c \to \infty,$$

i.e.,

$$\sup_{t \in [0,T]} \mathrm{Tr} |\gamma^{(1)}(t)| = 0.$$

Hence, we have finished the proof of the main theorem assuming Theorem 2.3. The bulk of the rest of the paper is devoted to proving Theorem 2.3.

3. Estimate for the nonlinear part

Recall

$$\mathrm{NP}^{(l_c)} = G^{(1)} + \sum_{r=1}^{l_c} \int_0^{t_1} \cdots \int_0^{t_r} dt_2 \cdots dt_{r+1} U^{(1)}(t_1 - t_2) B_2 \cdots U^{(r)}(t_r - t_{r+1}) B_{r+1} G^{(r+1)}(t_{r+1})$$

$$=: \mathrm{I} + \mathrm{II},$$

where

$$G^{(k)}(t_k) = \int_0^{t_k} dt_{k+1} U^{(k)}(t_k - t_{k+1}) \big(\mathcal{A}_1^{(k)} \gamma^{(k)}(t_{k+1}) + \mathcal{A}^{(k)} \gamma_2^{(k)}(t_{k+1}) \big). \tag{39}$$

We will first treat $\mathrm{Tr} |G^{(1)}(t_1)|$ coming from part I and then, with some additional tools, the corresponding term coming from part II. Both of the estimates rely upon the quantum de Finetti theorem stated below.

Theorem 3.1 (quantum de Finetti theorem [Hudson and Moody 1976; Størmer 1969; Ammari and Nier 2008; 2011; Lewin, Nam, and Rougerie 2014]). *Let $\mathcal{H}$ be a separable Hilbert space, and let $\mathcal{H}^k = \bigotimes_{\mathrm{sym}}^k \mathcal{H}$ denote the corresponding bosonic k-particle space. Let Γ denote a collection of bosonic density matrices on $\mathcal{H}$, i.e.,*

$$\Gamma = (\gamma^{(1)}, \gamma^{(2)}, \ldots)$$

with $\gamma^{(k)}$ a non-negative trace class operator on $\mathcal{H}^k$. If Γ is admissible, i.e., for all $k \in \mathbb{N}$ we have $\mathrm{Tr} \, \gamma^{(k)} = 1$ and $\gamma^{(k)} = \mathrm{Tr}_{k+1} \, \gamma^{(k+1)}$, where Tr_{k+1} denotes the partial trace over the $(k+1)$-th factor, then there exists a unique Borel probability measure μ, supported on the unit sphere in $\mathcal{H}$, and invariant under multiplication of $\phi \in \mathcal{H}$ by complex numbers of modulus one, such that

$$\gamma^{(k)} = \int d\mu(\phi) (|\phi\rangle\langle\phi|)^{\otimes k} \quad \text{for all } k \in \mathbb{N}.$$

Remark 3.2. The μ determined by Theorem 3.1 is finite and so, in particular, σ-finite. Therefore, the Fubini–Tonelli theorem, which is crucial in the proof, applies. See [Dunford and Schwartz 1988, p. 190].

Using Theorem 3.1, we write

$$\gamma_j^{(k)}(t) = \int d\mu_t^{(j)}(\phi)(|\phi\rangle\langle\phi|)^{\otimes k}, \quad j = 1, 2,$$

and

$$\gamma^{(k)}(t) = \int d\mu_t(\phi)(|\phi\rangle\langle\phi|)^{\otimes k},$$

where $\mu_t := \mu_t^{(1)} - \mu_t^{(2)}$ is a signed measure supported on the unit sphere of $L^2(\mathbb{R}^2)$. We remark that

$$\mathrm{Tr}|\gamma^{(1)}(t)| = \int d|\mu_t|(\phi)\|\phi\|_{L^2}^2 = \int d|\mu_t|(\phi)$$

while

$$\mathrm{Tr}|\gamma_j^{(1)}(t)| = \int d\mu_t^{(j)}\|\phi\|_{L^2}^2 = \int d\mu_t^{(j)} = 1.$$

Here $|\mu_t|$ is defined, in the usual way, as the sum of the positive part and the negative part of μ_t, which itself is another finite measure since $|\mu_t| \le \mu_t^{(1)} + \mu_t^{(2)}$. Write $\mu_t^{(0)} = \mu_t$ for convenience. The main properties of $\mu_t^{(i)}$ that we need are

$$\sup_{t \in [0,T]} \int d|\mu_t^{(i)}|(\phi)\|\phi\|_{H_x^{2/3}}^{2k} \le M^{2k} \quad \text{for } i = 0, 1, 2 \tag{40}$$

and

$$|\mu_t^{(i)}|\big(\big\{\phi \in L^2(\mathbb{R}^2) \mid \|\phi\|_{H^{2/3}} > M\big\}\big) = 0 \quad \text{for } i = 0, 1, 2, \tag{41}$$

where $|\mu_t^{(i)}|$ is of course $\mu_t^{(i)}$ if $i = 1$ or 2. For $i = 1, 2$, estimate (40) is equivalent to the energy condition

$$\sup_{t \in [0,T]} \mathrm{Tr}\bigg(\prod_{j=1}^k \langle\nabla_{x_j}\rangle^{2/3}\bigg)\gamma_i^{(k)}(t)\bigg(\prod_{j=1}^k \langle\nabla_{x_j'}\rangle^{2/3}\bigg) \le M^{2k} \quad \text{for } i = 1, 2, \tag{42}$$

and (41) then follows from (40) using Chebyshev's inequality.[3] The $i = 0$ case then follows from the definition.

Putting these structures into $\mathscr{A}$, for $\ell = 1, 2$, we have

$$\mathscr{A}_\ell^{(k)} f(t) = \iint d\mu_t^{(\ell)}(\psi)\, d\mu_t^{(\ell)}(\omega) \sum_{j=1}^k \big[a_{|\psi|^2,|\omega|^2}(r_j) - a_{|\psi|^2,|\omega|^2}(r_j')\big]f \tag{43}$$

and

$$\mathscr{A}^{(k)} f(t) = (\mathscr{A}_1^{(k)} - \mathscr{A}_2^{(k)})f$$

$$= \iint d\mu_t^{(1)}(\psi)\, d\mu_t(\omega) \sum_{j=1}^k \big[a_{|\psi|^2,|\omega|^2}(r_j) - a_{|\psi|^2,|\omega|^2}(r_j')\big]f$$

$$+ \iint d\mu_t(\psi)\, d\mu_t^{(2)}(\omega) \sum_{j=1}^k \big[a_{|\psi|^2,|\omega|^2}(r_j) - a_{|\psi|^2,|\omega|^2}(r_j')\big]f, \tag{44}$$

---

[3]See [T. Chen, Hainzl, Pavlović, and Seiringer 2014, Lemma 4.4] or [Hong, Taliaferro, and Xie 2014, (2.17)].

where $a_{|\psi|^2,|\omega|^2}$ is defined by

$$a_{|\psi|^2,|\omega|^2}(t,r) := A_0^{(|\psi|^2,|\omega|^2)}(t,r) + \frac{1}{r^2} A_\theta^{(|\psi|^2)}(t,r) A_\theta^{(|\omega|^2)}(t,r)$$

with

$$A_0^{(|\psi|^2,|\omega|^2)}(t,r) = -\int_r^\infty A_\theta^{(|\psi|^2)}(t,s)|\omega|^2(t,s)\,\frac{ds}{s}, \qquad A_\theta^{(\rho)}(t,r) = -\frac{1}{2}\int_0^r \rho(t,s)s\,ds.$$

Informally speaking, $a_{|\psi|^2,|\omega|^2}(r)$ is similar to $a(r)$ defined in (28) but is linear with respect to $|\psi|^2$ and $|\omega|^2$ independently rather than quadratic with respect to a single $|\phi|^2$.

This notation enables us to represent the core term of $G^{(k)}$ by

$$\mathcal{A}_1^{(k)} \gamma^{(k)}(t) + \mathcal{A}^{(k)} \gamma_2^{(k)}(t)$$
$$= \sum_{j=1}^k \sum_{(l,m,n)\in\mathcal{P}} \iiint d\mu_t^{(l)}(\psi)\,d\mu_t^{(m)}(\omega)\,d\mu_t^{(n)}(\phi) \big[a_{|\psi|^2,|\omega|^2}(r_j) - a_{|\psi|^2,|\omega|^2}(r_j')\big](|\phi\rangle\langle\phi|)^{\otimes k} \quad (45)$$

if we take $\mathcal{P} = \{(1,1,0),(2,2,0),(1,0,2),(0,2,2)\}$. The set $\mathcal{P}$ is for bookkeeping, incorporating the terms from (43) and (44), and we remind the readers that $d\mu_t^{(0)} := d\mu_t$. We remark that, to reach (45), we used the quantum de Finetti theorem (i.e., Theorem 3.1) four times: twice for the $\gamma^{(k)}$ term (once for γ_1 and once for γ_2) and twice for the terms in the self-generated potential $\mathcal{A}$ (they are quadratic in ρ).

3A. *Estimate of* $\mathrm{Tr}|G^{(1)}(t_1)|$. Putting $k=2$ in (45) and replacing ψ, ω, and ϕ with ϕ_1, ϕ_2, and ϕ_3, respectively, we have

$$\mathrm{Tr}|G^{(1)}(t_1)| = \mathrm{Tr}\left|\int_0^{t_1} dt_2\, U^{(1)}(t_1-t_2)\big(\mathcal{A}_1^{(1)}\gamma^{(1)}(t_2) + \mathcal{A}^{(1)}\gamma_2^{(1)}(t_2)\big)\right|$$
$$\leq \sum_{(l,m,n)\in\mathcal{P}} \int_0^{t_1} dt_2 \iiint d|\mu_{t_2}^{(l)}|(\phi_1)\,d|\mu_{t_2}^{(m)}|(\phi_2)\,d|\mu_{t_2}^{(n)}|(\phi_3)$$
$$\times \mathrm{Tr}\big|U^{(1)}(t_1-t_2)\big[a_{|\phi_1|^2,|\phi_2|^2}(r_1) - a_{|\phi_1|^2,|\phi_2|^2}(r_1')\big]\phi_3(r_1)\bar{\phi}_3(r_1')\big|.$$

Using the fact that

$$\mathrm{Tr}|U^{(1)}(t)f(r_1)g(r_1')| = \int \big|e^{it\Delta}f(r_1)e^{-it\Delta}g(r_1)\big|\,dx_1$$
$$\leq \|e^{it\Delta}f\|_{L_x^2}\|e^{it\Delta}g\|_{L_x^2}$$
$$= \|f\|_{L_x^2}\|g\|_{L_x^2},$$

we have

$$\mathrm{Tr}|G^{(1)}(t_1)| \leq \sum_{(l,m,n)\in\mathcal{P}} \int_0^{t_1} dt_2 \iiint d|\mu_{t_2}^{(l)}|(\phi_1)\,d|\mu_{t_2}^{(m)}|(\phi_2)\,d|\mu_{t_2}^{(n)}|(\phi_3)\|a_{|\phi_1|^2,|\phi_2|^2}\phi_3\|_{L_x^2}\|\phi_3\|_{L_x^2}.$$

Corollary 4.9, i.e., the main nonlinear estimate, turns the above into

$$\mathrm{Tr}|G^{(1)}(t_1)| \leq \sum_{(l,m,n)\in\mathcal{P}} \int_0^{t_1} dt_2 \iiint d|\mu_{t_2}^{(l)}|(\phi_1)\,d|\mu_{t_2}^{(m)}|(\phi_2)\,d|\mu_{t_2}^{(n)}|(\phi_3) \times \|\phi_3\|_{L_x^2}$$
$$\times \|\phi_1\|_{\dot{H}_x^{1/2}}\|\phi_2\|_{\dot{H}_x^{1/2}} \min_{\tau\in S_3}\|\phi_{\tau(1)}\|_{\dot{H}_x^{1/2}}\|\phi_{\tau(2)}\|_{\dot{H}_x^{1/2}}\|\phi_{\tau(3)}\|_{L_x^2}.$$

One of l, m, or n is zero, and we may put the corresponding term in L^2, i.e.,

$$\mathrm{Tr}|G^{(1)}(t_1)| \leq \sum_{l=1}^{2} \int_0^{t_1} dt_2 \iiint d\mu_{t_2}^{(l)}(\phi_1)\, d\mu_{t_2}^{(l)}(\phi_2)\, d|\mu_{t_2}^{(0)}|(\phi_3)\|\phi_1\|_{\dot{H}_x^{1/2}}^2\|\phi_2\|_{\dot{H}_x^{1/2}}^2\|\phi_3\|_{L_x^2}^2$$

$$+ \int_0^{t_1} dt_2 \iiint d\mu_{t_2}^{(1)}(\phi_1)\, d|\mu_{t_2}|(\phi_2)\, d\mu_{t_2}^{(2)}(\phi_3)\|\phi_1\|_{\dot{H}_x^{1/2}}^2\|\phi_2\|_{L_x^2}\|\phi_2\|_{\dot{H}_x^{1/2}}\|\phi_3\|_{\dot{H}_x^{1/2}}\|\phi_3\|_{L_x^2}$$

$$+ \int_0^{t_1} dt_2 \iiint d|\mu_{t_2}|(\phi_1)\, d\mu_{t_2}^{(2)}(\phi_2)\, d\mu_{t_2}^{(2)}(\phi_3)\|\phi_1\|_{L_x^2}\|\phi_1\|_{\dot{H}_x^{1/2}}\|\phi_2\|_{\dot{H}_x^{1/2}}^2\|\phi_3\|_{\dot{H}_x^{1/2}}\|\phi_3\|_{L_x^2}.$$

Using the fact that each $\mu_t^{(j)}$ is supported on the unit sphere in L^2 and thanks to (40) and (41), we obtain

$$\mathrm{Tr}|G^{(1)}(t_1)| \leq 4M^3 t_1 \sup_{t\in[0,t_1]} \int d|\mu_t|(\phi) \leq CM^3 t_1 \left(\sup_{t\in[0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| \right).$$

Thus, we have proved that

$$\mathrm{Tr}|G^{(1)}(t_1)| \leq Ct_1 \left(\sup_{t\in[0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| \right). \tag{46}$$

3B. *Estimate for part II.* Recall that

$$\mathrm{II} = \sum_{r=1}^{l_c} \int_0^{t_1} \cdots \int_0^{t_r} dt_2 \cdots dt_{r+1} U^{(1)}(t_1-t_2)B_2 \cdots U^{(r)}(t_r-t_{r+1})B_{r+1}G^{(r+1)}(t_{r+1}).$$

Because each B_j is a sum of $2(j-1)$ terms (see (31)), integrands of summands of $\mathrm{NP}^{(l_c)}$ have up to $O(k!)$ summands themselves. We use the Klainerman–Machedon board game argument to combine them and hence reduce the number of terms that need to be treated. Define

$$J(\underline{t}_{j+1})(f^{(j+1)}) = U^{(1)}(t_1-t_2)B_2 \cdots U^{(j)}(t_j-t_{j+1})B_{j+1}f^{(j+1)},$$

where $\underline{t}_{j+1}$ means $(t_2,\ldots,t_{j+1})$. Then the Klainerman–Machedon board game argument implies the lemma.

Lemma 3.3 (Klainerman–Machedon board game [2008]). *One can express*

$$\int_0^{t_1} \cdots \int_0^{t_j} J(\underline{t}_{j+1})(f^{(j+1)})\, d\underline{t}_{j+1}$$

as a sum of at most 4^j terms of the form

$$\int_D J(\underline{t}_{j+1}, \sigma)(f^{(j+1)})\, d\underline{t}_{j+1},$$

or in other words,

$$\int_0^{t_1} \cdots \int_0^{t_j} J(\underline{t}_{j+1})(f^{(j+1)})\, d\underline{t}_{j+1} = \sum_\sigma \int_D J(\underline{t}_{j+1}, \sigma)(f^{(j+1)})\, d\underline{t}_{j+1}.$$

Here $D \subset [0, t_2]^j$, the σ range over the set of maps from $\{2,\ldots,j+1\}$ to $\{1,\ldots,j\}$ satisfying $\sigma(2)=1$ and $\sigma(l) < l$ for all l, and

$$J(\underline{t}_{j+1}, \sigma)(f^{(j+1)}) = U^{(1)}(t_1-t_2)B_{1,2}U^{(2)}(t_2-t_3)B_{\sigma(3),3} \cdots U^{(j)}(t_j-t_{j+1})B_{\sigma(j+1),j+1}(f^{(j+1)}).$$

With Lemma 3.3, we can write a typical summand of part II as

$$\int_0^{t_1} \cdots \int_0^{t_r} dt_2 \cdots dt_{r+1} U^{(1)}(t_1 - t_2) B_2 \cdots U^{(r)}(t_r - t_{r+1}) B_{r+1} G^{(r+1)}(t_{r+1})$$
$$= \sum_\sigma \int_D d\underline{t}_{r+1} J(\underline{t}_{r+1}, \sigma)(G^{(r+1)}),$$

where the sum has at most 4^r terms inside. Let

$$\text{II}^{(r,\sigma)} = \int_D d\underline{t}_{r+1} J(\underline{t}_{r+1}, \sigma)(G^{(r+1)}). \tag{47}$$

To estimate part II, it suffices to prove the following lemma:

Lemma 3.4. *There is a C_0 depending on M in (42) such that, for all r, we have*

$$\text{Tr}|\text{II}^{(r,\sigma)}|(t_1) \leq \left[(r+1)\left(C_0 t_1^{1/3}\right)^r\right] t_1 \left(\sup_{t \in [0,t_1]} \text{Tr}|\gamma^{(1)}(t)|\right).$$

With the above lemma, we have

$$\text{Tr}|\text{II}|(t_1) \leq \sum_{r=1}^{l_c} \sum_\sigma \left[(r+1)\left(C_0 t_1^{1/3}\right)^r\right] t_1 \left(\sup_{t \in [0,t_1]} \text{Tr}|\gamma^{(1)}(t)|\right)$$
$$\leq t_1 \left(\sup_{t \in [0,t_1]} \text{Tr}|\gamma^{(1)}(t)|\right) \sum_{r=1}^\infty 4^r \left[(r+1)\left(C_0 t_1^{1/3}\right)^r\right]$$
$$\leq C t_1 \left(\sup_{t \in [0,t_1]} \text{Tr}|\gamma^{(1)}(t)|\right) \tag{48}$$

for t_1 small enough so that the series converges.

Together the estimates (46) and (48) establish Theorem 2.3.

Before proving Lemma 3.4, we illustrate how to obtain the estimate for a specific example.

Example 3.5. To avoid heavy notation and demonstrate the main idea of the proof of Lemma 3.4, we first prove it for a concrete example. The general case uses the same underlying idea, which turns out to be quite simple as compared to what must be done for the interaction part IP. We adapt the example and use the notation in [T. Chen, Hainzl, Pavlović, and Seiringer 2014, §6.1] for our $\text{II}^{(r,\sigma)}$. Denoting $U^{(j)}(t_k - t_l)$ by $U^{(j)}_{k,l}$, we consider

$$\text{Tr}|\text{II}^{(3,\sigma)}|(t_1) = \int_D d\underline{t}_4 U^{(1)}_{1,2} B_{1,2} U^{(2)}_{2,3} B_{2,3} U^{(3)}_{3,4} B_{3,4} [G^{(4)}(t_4)]$$

$$\leq \sum_{j=1}^4 \sum_{(l,m,n) \in \mathscr{P}} \int_{[0,t_1]^3} d\underline{t}_4 \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi) \, d|\mu_{t_5}^{(m)}|(\omega) \, d|\mu_{t_5}^{(n)}|(\phi)$$
$$\times \text{Tr}\left|U^{(1)}_{1,2} B_{1,2} U^{(2)}_{2,3} B_{2,3} U^{(3)}_{3,4} B_{3,4} U^{(4)}_{4,5}\left([a_{|\psi|^2,|\omega|^2}(r_j) - a_{|\psi|^2,|\omega|^2}(r_j')](|\phi\rangle\langle\phi|)^{\otimes 4}\right)\right|. \tag{49}$$

Remark 3.6. In the above, there is a $U^{(4)}_{4,5}$ after $B_{3,4}$. This is the main difference between the nonlinear part NP and the interaction part IP. As noted in [T. Chen, Hainzl, Pavlović, and Seiringer 2014], since

the last B in IP is not followed by a Schrödinger propagator, it creates a factor of $|\phi|^2\phi$, which has to be handled by Sobolev embedding rather than Strichartz estimates.

It suffices to treat

$$\sum_{(l,m,n)\in\mathscr{P}} \int_{[0,t_1]^3} dt_4 \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi)$$
$$\times \mathrm{Tr}\big|U_{1,2}^{(1)}B_{1,2}^+ U_{2,3}^{(2)}B_{2,3}^+ U_{3,4}^{(3)}B_{3,4}^+ U_{4,5}^{(4)}\big([a_{|\psi|^2,|\omega|^2}(r_4)](|\phi\rangle\langle\phi|)^{\otimes 4}\big)\big|, \quad (50)$$

where $B_{1,2}^+$ is half of $B_{1,2}$, namely

$$B_{1,2}^+(\gamma^{(2)}) = \gamma^{(2)}(x_1, x_1, x_1', x_1).$$

When we plug the estimate of (50) into (49), we will pick up a 2^3 since there are three B's in (49). However, compensating for this is the factor $\big(t_1^{2/3}\big)^3$ that emerges by the end. Hence, our simplification is a valid one.

Step I (structure). We enumerate the four factors of $(|\phi\rangle\langle\phi|)^{\otimes 4}$ for the purpose of bookkeeping even though these factors are physically indistinguishable. So we write $\bigotimes_{i=1}^4 u_i$, ordered with increasing index i. We first have

$$B_{3,4}^+ U_{4,5}^{(4)} a_{|\psi|^2,|\omega|^2}(r_4)(|\phi\rangle\langle\phi|)^{\otimes 4} = \left(U_{4,5}^{(2)}\left(\bigotimes_{i=1}^2 u_i\right)\right)\otimes\Theta_3,$$

where

$$\begin{aligned}
\Theta_3 &= B_{1,2}^+(U_{4,5}^{(2)}(u_3 \otimes a_{|\psi|^2,|\omega|^2}(r_4)u_4))\\
&= B_{1,2}^+(U_{4,5}\phi(x_3))(U_{5,4}\bar\phi(x_3'))(U_{4,5}[a_{|\psi|^2,|\omega|^2}(r_4)\phi(x_4)])(U_{5,4}\bar\phi(x_4'))\\
&= (U_{4,5}\phi(x_3))(U_{4,5}[a_{|\psi|^2,|\omega|^2}(r_3)\phi(x_3)])(U_{5,4}\bar\phi(x_3))(U_{5,4}\bar\phi(x_3'))\\
&\equiv T_3(x_3)(U_{5,4}\bar\phi(x_3')) \qquad\qquad (51)
\end{aligned}$$

with $U_{4,5} = e^{i(t_4-t_5)\Delta}$. Here T_3 stands for the trilinear form

$$(U_{4,5}[a_{|\psi|^2,|\omega|^2}(r_3)\phi(x_3)])(U_{5,4}\bar\phi(x_3))(U_{5,4}\bar\phi(x_3')).$$

We make similar substitutions below and, to bound these terms, shall invoke the trilinear estimate (56), which states that

$$\|T(f_1, f_2, f_3)\|_{L^1_{t\in[0,t_0)}L^2_x} \lesssim_s t_0^{s/2}\|f_1\|_{L^2}\|f_2\|_{L^2}\|f_3\|_{\dot H^s}$$

for $0 < s \leq 2$.

Applying $B_{2,3}^+ U_{3,4}^{(3)}$, we reach

$$\begin{aligned}
B_{2,3}^+ U_{3,4}^{(3)} B_{3,4}^+ U_{4,5}^{(4)}(a_{\rho_l,\rho_m}(r_j)(|\phi\rangle\langle\phi|)^{\otimes 4}) &= B_{2,3}^+ U_{3,4}^{(3)}(U_{4,5}^{(1)}u_1 \otimes U_{4,5}^{(1)}u_2 \otimes \Theta_3)\\
&= U_{3,4}^{(1)} U_{4,5}^{(1)}u_1 \otimes \Theta_2\\
&= U_{3,5}^{(1)}u_1 \otimes \Theta_2,
\end{aligned}$$

where

$$
\begin{aligned}
\Theta_2 &= B_{1,2}^+ U_{3,4}^{(2)}(U_{4,5}^{(1)} u_2 \otimes \Theta_3) \\
&= B_{1,2}^+ (U_{3,5}^{(1)} u_2 \otimes U_{3,4}^{(1)} \Theta_3) \\
&= B_{1,2}^+ \big((U_{3,5}\phi(x_2))(U_{5,3}\bar{\phi}(x_2'))(U_{3,4}T_3(x_3))(U_{4,3}U_{5,4}\bar{\phi}(x_3')) \big) \\
&= (U_{3,5}\phi(x_2))(U_{3,4}T_3(x_2))(U_{5,3}\bar{\phi}(x_2))(U_{5,3}\bar{\phi}(x_2')) \\
&\equiv T_2(x_2)(U_{5,3}\bar{\phi}(x_2')).
\end{aligned}
\tag{52}
$$

Finally, with $U_{1,2}^{(1)} B_{1,2}^+ U_{2,3}^{(2)}$, we get

$$
\begin{aligned}
U_{1,2}^{(1)} B_{1,2}^+ U_{2,3}^{(2)} B_{2,3}^+ U_{3,4}^{(3)} B_{3,4}^+ U_{4,5}^{(4)} & (a_{|\psi|^2,|\omega|^2}(r_j)(|\phi\rangle\langle\phi|)^{\otimes 4}) \\
&= U_{1,2}^{(1)} B_{1,2}^+ U_{2,3}^{(2)}(U_{3,5}^{(1)} u_1 \otimes \Theta_2) \\
&= U_{1,2}^{(1)} B_{1,2}^+ (U_{2,5}^{(1)} u_1 \otimes U_{2,3}^{(1)} \Theta_2) \\
&= U_{1,2}^{(1)} B_{1,2}^+ \big[(U_{2,5}\phi(x_1))(U_{5,2}\bar{\phi}(x_1'))(U_{2,3}T_2(x_2))(U_{3,2}U_{5,3}\bar{\phi}(x_2')) \big] \\
&= U_{1,2}^{(1)} \big[(U_{2,5}\phi(x_1))(U_{2,3}T_2(x_1))(U_{5,2}\bar{\phi}(x_1))(U_{5,2}\bar{\phi}(x_1')) \big] \\
&= U_{1,2}^{(1)} [T_1(x_1)U_{5,2}\bar{\phi}(x_1')].
\end{aligned}
\tag{53}
$$

Step II (iterative estimate). Plugging the calculation in Step I into (50), we have

$$
\begin{aligned}
(50) &\leq \sum_{(l,m,n)\in\mathscr{P}} \int_{[0,t_1]^3} dt_4 \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi) \|T_1(x_1)\|_{L^2} \|\bar{\phi}\|_{L^2} \\
&\leq \sum_{(l,m,n)\in\mathscr{P}} \int_{[0,t_1]^2} dt_3 t_4 \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi) \times \|T_1\|_{L_{t_2}^1 L^2},
\end{aligned}
$$

where

$$
\|T_1\|_{L_{t_2}^1 L^2} \leq Ct_1^{1/3} \|\phi\|_{H^{2/3}} \|T_2\|_{L^2} \|\phi\|_{L^2}
$$

by (56). Thus,

$$
(50) \leq Ct_1^{1/3} \sum_{(l,m,n)\in\mathscr{P}} \int_{[0,t_1]} t_4 \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi) \|\phi\|_{H^{2/3}} \|T_2\|_{L_{t_3}^1 L^2}.
$$

By (56) again,

$$
\|T_2(x_2)\|_{L_{t_3}^1 L_{x_2}^2} \leq Ct_1^{1/3} \|\phi\|_{H^{2/3}} \|T_3\|_{L^2} \|\phi\|_{L^2},
$$

and hence,

$$
\begin{aligned}
(50) &\leq (Ct_1^{1/3})^2 \sum_{(l,m,n)\in\mathscr{P}} \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi) \|\phi\|_{H^{2/3}}^2 \|T_3\|_{L_{t_4}^1 L_{x_2}^2} \\
&\leq (Ct_1^{1/3})^3 \sum_{(l,m,n)\in\mathscr{P}} \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi)\, d|\mu_{t_5}^{(m)}|(\omega)\, d|\mu_{t_5}^{(n)}|(\phi) \|\phi\|_{H^{2/3}}^3 \|a_{|\psi|^2,|\omega|^2}(r_3)\phi(x_3)\|_{L^2}.
\end{aligned}
$$

By the fact that $|\mu_t^{(i)}|$ is supported in the set

$$\{\phi \in L^2(\mathbb{R}^2) \mid \|\phi\|_{H^{2/3}} \le M\},$$

we have

$$(50) \le \left(CMt_1^{1/3}\right)^3 \sum_{(l,m,n)\in\mathscr{P}} \int_0^{t_1} dt_5 \iiint d|\mu_{t_5}^{(l)}|(\psi) \, d|\mu_{t_5}^{(m)}|(\omega) \, d|\mu_{t_5}^{(n)}|(\phi) \left\| a_{|\psi|^2, |\omega|^2}(r_3)\phi(x_3) \right\|_{L^2}.$$

One then proceeds as in the estimate of $\mathrm{Tr}|G^{(1)}(t_1)|$ to reach

$$(50) \le \left(CMt_1^{1/3}\right)^3 M^3 t_1 \left(\sup_{t\in[0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| \right).$$

Selecting a C_0 bigger than M^2 and 1, we obtain

$$(50) \le \left(C_0 t_1^{1/3}\right)^3 t_1 \left(\sup_{t\in[0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| \right).$$

Plugging the above estimate back into (49), we get

$$\mathrm{Tr}|\mathrm{II}^{(3,\sigma)}|(t_1) \le \left[4\cdot 2^3 \cdot \left(C_0 t_1^{1/3}\right)^3\right] t_1 \left(\sup_{t\in[0,t_1]} \mathrm{Tr}|\gamma^{(1)}(t)| \right)$$

as desired. This finishes the proof of the example.

One observation to make concerning our approach in Example 3.5 is that the structure found in Step I is crucial. Such a structure generated by the collision operator B and propagator U is found in general, and we state its relevant properties in the following lemma:

Lemma 3.7. *Let $M \in \mathbb{N}$, $M > 1$, and for each j, $1 \le j \le M$, suppose that the two functions $f_j(x_j)$ and $f_j'(x_j')$ belong to*[4] *$L_t^1 H_x^s(\mathbb{R}^2)$, $\frac{1}{2} \le s \le \frac{2}{3}$. Then there exist $L_t^1 H_x^s(\mathbb{R}^2)$ functions h and h' such that*

$$B_{\sigma(M),M}^{\pm} U_{M,M+1}^{(M)} \left[\prod_{j=1}^{M} f_j(x_j) f_j'(x_j') \right] = h_{\sigma(M)}(x_{\sigma(M)}) h_{\sigma(M)}'(x_{\sigma(M)}') U_{M,M+1}^{(M-2)} \left[\prod_{\substack{j=1 \\ j\ne\sigma(M)}}^{M-1} f_j(x_j) f_j'(x_j') \right].$$

In the case where B is $B_{\sigma(M),M}^+$, h is a trilinear form of the type (54) and h' is a linear evolution. In the case where B is $B_{\sigma(M),M}^-$, the roles of h and h' are reversed.

Proof. The collision operator leaves untouched each term for which $j \notin \{M, \sigma(M)\}$. Only the propagator affects these terms. So we have

$$B_{\sigma(M),M}^{+} U_{M,M+1}^{(M)} \left[\prod_{j=1}^{M} f_j(x_j) f_j'(x_j') \right]$$

$$= U_{M,M+1}^{(M-2)} \left[\prod_{j\in\{1,\dots,M\}\setminus\{M,\sigma(M)\}} f_j(x_j) f_j'(x_j') \right] \cdot T_{\sigma(M),M}(x_{\sigma(M)}) e^{-i(t_M - t_{M+1})\Delta_{x_{\sigma(M)}'}} f_{\sigma(M)}'(x_{\sigma(M)}'),$$

---

[4] We suppress the time dependence in the notation and allow restriction to time intervals, which may be achieved, for instance, by introducing sharp time cutoffs.

where

$$T_{\sigma(M),M}(x_{\sigma(M)})$$
$$:= e^{i(t_M - t_{M+1})\Delta_{x_{\sigma(M)}}} f_{\sigma(M)}(x_{\sigma(M)}) \cdot e^{i(t_M - t_{M+1})\Delta_{x_{\sigma(M)}}} f_M(x_{\sigma(M)}) \cdot e^{-i(t_M - t_{M+1})\Delta_{x_{\sigma(M)}}} f'_M(x_{\sigma(M)}).$$

Similarly,

$$B^-_{\sigma(M),M} U^{(M)}_{M,M+1} \left[\prod_{j=1}^{M} f_j(x_j) f'_j(x'_j) \right]$$
$$= U^{(M-2)}_{M,M+1} \left[\prod_{j \in \{1,\ldots,M\}\setminus\{M,\sigma(M)\}} f_j(x_j) f'_j(x'_j) \right] \cdot T'_{\sigma(M),M}(x'_{\sigma(M)}) e^{i(t_M - t_{M+1})\Delta_{x_{\sigma(M)}}} f_{\sigma(M)}(x_{\sigma(M)}),$$

where

$$T'_{\sigma(M),M}(x'_{\sigma(M)})$$
$$:= e^{i(t_M - t_{M+1})\Delta_{x'_{\sigma(M)}}} f_M(x'_{\sigma(M)}) \cdot e^{-i(t_M - t_{M+1})\Delta_{x'_{\sigma(M)}}} f'_{\sigma(M)}(x'_{\sigma(M)}) \cdot e^{-i(t_M - t_{M+1})\Delta_{x'_{\sigma(M)}}} f'_M(x'_{\sigma(M)}).$$

The $L^1_t H^s_x$ bounds follow from (55) and Strichartz. $\qquad\square$

Proof of Lemma 3.4. Using (47), (39), and (45), we write

$$\mathrm{II}^{(r,\sigma)} = \sum_{j=1}^{r+1} \sum_{(l,m,n)\in\mathscr{P}} \int_D d\underline{t}_{r+1} J(\underline{t}_{r+1}, \sigma) \left\{ \int_0^{t_{r+1}} dt_{r+2} U^{(r+1)}(t_{r+1} - t_{r+2}) \right.$$
$$\left. \times \iiint d\mu^{(l)}_{t_{r+2}}(\psi) \, d\mu^{(m)}_{t_{r+2}}(\omega) \, d\mu^{(n)}_{t_{r+2}}(\phi) [a_{|\psi|^2,|\omega|^2}(|x_j|) - a_{|\psi|^2,|\omega|^2}(|x'_j|)](|\phi\rangle\langle\phi|)^{\otimes(r+1)} \right\}.$$

We abbreviate

$$J(\underline{t}_{r+1}, \sigma) = U^{(1)}_{1,2} B_{1,2} U^{(2)}_{2,3} B_{\sigma(3),3} \cdots U^{(r)}_{r,r+1} B_{\sigma(r+1),r+1}$$

and write

$$\mathrm{Tr}|\mathrm{II}^{(r,\sigma)}|(t_1)$$
$$\leq \sum_{j=1}^{r+1} \sum_{(l,m,n)\in\mathscr{P}} \int_{[0,t_1]^r} d\underline{t}_{r+1} \int_0^{t_1} dt_{r+2} \iiint d|\mu^{(l)}_{t_{r+2}}|(\psi) \, d|\mu^{(m)}_{t_{r+2}}|(\omega) \, d|\mu^{(n)}_{t_{r+2}}|(\phi)$$
$$\times \mathrm{Tr}\big|U^{(1)}_{1,2} B_{1,2} \cdots U^{(r)}_{r,r+1} B_{\sigma(r+1),r+1} U^{(r+1)}_{r+1,r+2} [a_{|\psi|^2,|\omega|^2}(|x_j|) - a_{|\psi|^2,|\omega|^2}(|x'_j|)](|\phi\rangle\langle\phi|)^{\otimes(r+1)}\big|.$$

To simplify calculations, we drop, without loss of generality, the $-a_{|\psi|^2,|\omega|^2}(|x'_j|)$ term. Also, we split each $B_{j,k}$ into two pieces $B^{\pm}_{j,k}$ so that $B_{j,k} = B^+_{j,k} - B^-_{j,k}$.

Consider first the innermost terms

$$B^{\pm}_{\sigma(r+1),r+1} U^{(r+1)}_{r+1,r+2} a_{|\psi|^2,|\omega|^2}(|x_j|)(|\phi\rangle\langle\phi|)^{\otimes(r+1)}.$$

The index $j \in \{1, \ldots, r+1\}$ and the permutation σ together determine at what point $a_{|\psi|^2,|\omega|^2}(|x_j|)$ is directly affected by a collision operator. In any case, we claim that, with respect to the variables $x_{\sigma(r+1)}$

and $x'_{\sigma(r+1)}$, the term

$$B^+_{\sigma(r+1),r+1} U^{(r+1)}_{r+1,r+2} a_{|\psi|^2,|\omega|^2}(|x_j|)(|\phi\rangle\langle\phi|)^{\otimes(r+1)}$$

is a trilinear form of the form T in (54) (see (51), (52), and (53) for examples of these trilinear forms) in the $x_{\sigma(r+1)}$ variable and a linear flow in the $x'_{\sigma(r+1)}$ variable (the term with B^- instead of B^+ is similar but with the roles of the primed and unprimed variables reversed). Note that precisely one of the terms in the trilinear form T involves $a_{|\psi|^2,|\omega|^2}(|x_j|)$. This follows from Lemma 3.7. Additionally, Lemma 3.7 is formulated so that we can apply it iteratively until termination, at which point we have one term that is trilinear of the form (54) in precisely one of x_1 or x'_1 and another term that is a linear evolution of a function of the remaining spatial variable. Step I of Example 3.5 illustrates such a process.

The final step is to iteratively bound the terms. We follow Step II of Example 3.5. The underlying idea behind the iterative bounds is relatively straightforward. We start by controlling the trace norm using Cauchy–Schwarz in space. One factor is simply a ϕ term associated to the measure and so will have L^2 norm equal to one. This leaves us with the other term in $L^1_t L^2$. The next step is to apply (56). This places one factor in $\dot{H}^s$ and the remaining ones in L^2. So that we can eventually apply (70), it is important to always place in L^2 the term appearing in the right-hand side that involves $a_{|\psi|^2,|\omega|^2}(|x_j|)$. To control the term placed in $\dot{H}^s$, we apply (55). For the terms in L^2, we use (56) or (70) as appropriate. $\square$

Remark 3.8. We first remind the reader that, because at each step we are estimating a linear term of the type $e^{it\Delta} f$ or a trilinear term of the form (54), we do not need to apply Sobolev embedding as is necessary for estimating the interaction part. Secondly, the "a" term cannot be generated by B, and thus, we do not need to keep track of multiple "copies" of $|\phi|^2\phi$ generated by B in contrast to what must be done in controlling the interaction part. In particular, there is no need to introduce binary tree graphs or keep track of complicated factorization structures of kernels in controlling the nonlinear part.

4. Multilinear estimates

In this section, we will need the following fractional Leibniz rule from [Christ and Weinstein 1991, Proposition 3.3]:

Lemma 4.1. *Let $0 < s \leq 1$ and $1 < r, p_1, p_2, q_1, q_2 < \infty$ such that $\frac{1}{r} = \frac{1}{p_i} + \frac{1}{q_i}$ for $i = 1, 2$. Then*

$$\||\nabla|^s (fg)\|_{L^r} \lesssim \|f\|_{L^{p_1}} \||\nabla|^s g\|_{L^{q_1}} + \||\nabla|^s f\|_{L^{p_2}} \|g\|_{L^{q_2}}.$$

Define the trilinear form T by

$$T(f, g, h) = e^{i(t-t_1)\Delta} f \cdot e^{i(t-t_2)\Delta} g \cdot e^{i(t-t_3)\Delta} h. \tag{54}$$

Lemma 4.2. *Let $0 < s \leq \frac{2}{3}$. The trilinear form T given by (54) satisfies*[5]

$$\|T(f, g, h)\|_{L^1_{t\in[0,t_0)} \dot{H}^s_x} \lesssim t_0^s \|f\|_{\dot{H}^s} \|g\|_{\dot{H}^s} \|h\|_{\dot{H}^s}. \tag{55}$$

---

[5]Such trilinear estimates are the precursors to the Klainerman–Machedon collapsing estimates widely used in the literature. For those estimates, see [Klainerman and Machedon 2008; Kirkpatrick, Schlein, and Staffilani 2011; Grillakis and Margetis 2008; T. Chen and Pavlović 2011; X. Chen 2011; 2012a; Beckner 2014; Gressman, Sohinger, and Staffilani 2014].

Proof. By the fractional Leibniz rule, we have

$$\|T(f,g,h)\|_{L_t^1 \dot{H}_x^s} \lesssim \|e^{i(t-t_1)\Delta}f\|_{L_t^3 \dot{W}_x^{s,6}}\|e^{i(t-t_2)\Delta}g\|_{L_t^3 L_x^6}\|e^{i(t-t_3)\Delta}h\|_{L_t^3 L_x^6}$$

$$+ \|e^{i(t-t_1)\Delta}f\|_{L_t^3 L_x^6}\|e^{i(t-t_2)\Delta}g\|_{L_t^3 \dot{W}_x^{s,6}}\|e^{i(t-t_3)\Delta}h\|_{L_t^3 L_x^6}$$

$$+ \|e^{i(t-t_1)\Delta}f\|_{L_t^3 L_x^6}\|e^{i(t-t_2)\Delta}g\|_{L_t^3 L_x^6}\|e^{i(t-t_3)\Delta}h\|_{L_t^3 \dot{W}_x^{s,6}}.$$

By Sobolev embedding, we bound the first term by

$$\|e^{i(t-t_1)\Delta}f\|_{L_t^3 \dot{W}_x^{s,6}}\|e^{i(t-t_2)\Delta}g\|_{L_t^3 \dot{W}_x^{s,p}}\|e^{i(t-t_3)\Delta}h\|_{L_t^3 \dot{W}_x^{s,p}},$$

where $\frac{1}{p} = \frac{1}{6} + \frac{s}{2}$. Note that $2 \leq p < 6$. Let q be given by $\frac{1}{q} + \frac{1}{p} = \frac{1}{2}$ so that (q, p) forms a Schrödinger-admissible Strichartz pair (see, for instance, [Tao 2006, §2]). So we use Hölder in time to bound the expression by

$$\|e^{i(t-t_1)\Delta}f\|_{L_t^3 \dot{W}_x^{s,6}} t_0^{1/3-1/q}\|e^{i(t-t_2)\Delta}g\|_{L_t^q \dot{W}_x^{s,p}} t_0^{1/3-1/q}\|e^{i(t-t_3)\Delta}h\|_{L_t^q \dot{W}_x^{s,p}}.$$

Finally, we conclude by applying Strichartz estimates and noting that $\frac{1}{3} - \frac{1}{q} = \frac{s}{2}$. The second and third terms are similar. $\qquad\square$

Lemma 4.3. *Let $0 < s \leq 2$. The trilinear form T given by* (54) *satisfies*

$$\|T(f,g,h)\|_{L_{t\in[0,t_0)}^1 L_x^2} \lesssim t_0^{s/2}\|f\|_{L^2}\|g\|_{L^2}\|h\|_{\dot{H}^s}. \tag{56}$$

Proof. By Hölder's inequality,

$$\|T(f,g,h)\|_{L_{t\in[0,t_0)}^1 L_x^2} \leq t_0^{s/2}\|e^{i(t-t_1)\Delta}f\|_{L_t^q L_x^r}\|e^{i(t-t_2)\Delta}g\|_{L_t^q L_x^r}\|e^{i(t-t_3)\Delta}h\|_{L_t^\infty L_x^p},$$

where $\frac{1}{q} = \frac{1}{2} - \frac{s}{4}$, $r = \frac{4}{s}$, and $p = 2/(1-s)$. Using Strichartz estimates and Sobolev embedding, we control the right-hand side by

$$t_0^{s/2}\|f\|_{L_x^2}\|g\|_{L_x^2}\|e^{i(t-t_3)\Delta}h\|_{L_t^\infty \dot{H}_x^s}.$$

Finally, we conclude the bound stated in the lemma by noting that the Schrödinger propagator is an isometry on L^2-based spaces. $\qquad\square$

Remark 4.4. From the proofs of both (55) and (56), it is evident that any of $e^{i(t-t_1)\Delta}f$, $e^{i(t-t_2)\Delta}g$, and $e^{i(t-t_3)\Delta}h$ can be replaced by its complex conjugate in the trilinear form (54).

For the next set of estimates, recall

$$\partial_r A_0 = \frac{1}{r}A_\theta \rho, \qquad \partial_r A_\theta = -\frac{1}{2}r\rho$$

and

$$A_0(t,r) := -\int_r^\infty \frac{A_\theta(s)}{s}\rho(s)\,ds, \qquad A_\theta(t,r) := -\frac{1}{2}\int_0^r \rho(s)s\,ds. \tag{57}$$

When it is important to indicate the dependence upon the density function ρ, we write $A_\theta^{(\rho)}(t,r)$ for $A_\theta(t,r)$. Recall

$$A_0^{(\rho_1,\rho_2)}(t,r) = -\int_r^\infty A_\theta^{(\rho_1)}(s)\rho_2(s)\,\frac{ds}{s}, \tag{58}$$

where $A_\theta^{(\rho_1)}$ is defined using (57) but with ρ_1 in place of ρ in the right-hand side, i.e.,

$$A_\theta^{(\rho_1)}(t, r) = -\frac{1}{2} \int_0^r \rho_1(s) s \, ds.$$

Define the operators $[\partial_r]^{-1}$, $[r^{-n}\bar{\partial}_r]^{-1}$, and $[r\partial_r]^{-1}$ acting on radial functions by

$$[\partial_r]^{-1} f(r) = -\int_r^\infty f(s) \, ds, \qquad [r^{-n}\bar{\partial}_r]^{-1} f(r) = \int_0^r f(s) s^n \, ds, \qquad [r\partial_r]^{-1} f(r) = -\int_r^\infty \frac{1}{s} f(s) \, ds.$$

Then it follows by a direct argument that

$$\|[r\partial_r]^{-1} f\|_{L^p} \lesssim_p \|f\|_{L^p}, \quad 1 \le p < \infty, \tag{59}$$

$$\|r^{-n-1}[r^{-n}\bar{\partial}_r]^{-1} f\|_{L^p} \lesssim_p \|f\|_{L^p}, \quad 1 < p \le \infty, \tag{60}$$

$$\|[\partial_r]^{-1} f\|_{L^2} \lesssim \|f\|_{L^1}. \tag{61}$$

These estimates appear, for instance, in [Bejenaru, Ionescu, Kenig, and Tataru 2013, (1.5)] and also find application in [Liu and Smith 2014, §2].

Remark 4.5. In these estimates and those below, we use the Lebesgue measure on $\mathbb{R}^2$ for all L^p spaces. In particular, for radial functions of r, we essentially adopt the $r \, dr$ measure.

Lemma 4.6 (elementary bounds for A). *The connection coefficients A_θ and A_0, given by* (57), *satisfy*

$$\|A_\theta\|_{L_x^\infty} \lesssim \|\rho\|_{L_x^1}, \qquad \left\|\frac{1}{r} A_\theta\right\|_{L_x^\infty} \lesssim \|\rho\|_{L_x^2}, \qquad \left\|\frac{1}{r^2} A_\theta\right\|_{L_x^p} \lesssim \|\rho\|_{L_x^p}, \quad \text{where } 1 < p \le \infty, \tag{62}$$

and

$$\|A_0\|_{L_x^p} \lesssim \|\rho\|_{L_x^1} \|\rho\|_{L_x^p}, \quad \text{where } 1 \le p < \infty, \qquad \|A_0\|_{L_x^\infty} \lesssim \|\rho\|_{L_x^2}^2. \tag{63}$$

Moreover, A_θ^2 satisfies the bounds

$$\left\|\frac{1}{r^2} A_\theta^2\right\|_{L_x^p} \lesssim \|\rho\|_{L_x^1} \|\rho\|_{L_x^p}, \quad \text{where } 1 < p \le \infty, \qquad \left\|\frac{1}{r^2} A_\theta^2\right\|_{L_x^\infty} \lesssim \|\rho\|_{L_x^2}^2. \tag{64}$$

Proof. These estimates are essentially contained in [Liu and Smith 2014, §2].

The first inequality of (62) is trivial. The second follows from Cauchy–Schwarz:

$$|A_\theta(t, r)| \lesssim r \left(\int_0^\infty |\rho(s)|^2 s \, ds\right)^{1/2}.$$

The third is an application of (60) with $n = 1$.

The first inequality of (63) follows from the first inequality of (62) and from (59). The second is a consequence of Cauchy–Schwarz and the third inequality of (62) with $p = 2$.

The first inequality of (64) follows from the first and third inequalities of (62). The second follows from two applications of the second inequality of (62). $\qquad\square$

Lemma 4.7 (weighted estimates). *Let $\frac{1}{p} + \frac{1}{q} = 1$ with $1 < q < \infty$, and suppose that $\rho = |\psi|^2$ and $\rho_j = |\psi_j|^2$ for $j = 1, 2$. Then*

$$\|r^{-2/q} A_\theta^{(\rho)}\|_{L_x^\infty} \lesssim \|\psi\|_{\dot{H}_x^{1/q}}^2, \tag{65}$$

$$\|r^{-1/q} A_\theta^{(\rho)}\|_{L_x^\infty} \lesssim \|\psi\|_{\dot{H}_x^{1/q}} \|\psi\|_{L_x^2} \tag{66}$$

and

$$\|r^{1/p} A_0^{(\rho_1,\rho_2)}\|_{L_x^\infty} \lesssim \min_{\tau \in S_2} \|\psi_{\tau(1)}\|_{\dot{H}_x^{1/q}}^2 \|\psi_{\tau(2)}\|_{\dot{H}_x^{1/p}} \|\psi_{\tau(2)}\|_{L_x^2}, \tag{67}$$

where S_2 denotes the set of permutations on two elements.

Proof. To establish (66), use Hölder's inequality to obtain

$$|A_\theta| \lesssim r^{2/q} \|\psi\|_{L^{2p}}^2$$

and then use Sobolev embedding. The estimate (65) follows from Hölder's inequality, which yields

$$|A_\theta| \lesssim r^{1/q} \|r^{-1/q} \psi\|_{L_x^2} \|\psi\|_{L_x^2},$$

and Hardy's inequality.

To prove (67), use Hölder to write

$$|A_0^{(\rho_1,\rho_2)}| \lesssim \|r^{-2/q} A_\theta^{(\rho_1)}\|_{L_x^\infty} \|r^{-1/p} \psi_2\|_{L_x^2} \|\psi_2\|_{L_x^2} r^{-1/p}.$$

Then, using (65) and Hardy's inequality, we obtain

$$\|r^{1/p} A_0^{(\rho_1,\rho_2)}\|_{L_x^\infty} \lesssim \|\psi_1\|_{\dot{H}_x^{1/q}}^2 \|\psi_2\|_{\dot{H}_x^{1/p}} \|\psi_2\|_{L_x^2}.$$

Finally, we may repeat the argument with the roles of ψ_1 and ψ_2 reversed. $\qquad\square$

Lemma 4.8 (bounds for the nonlinear terms). *Suppose that $\rho_j = |\psi_j|^2$ for $j = 1, 2$. Then*

$$\|A_0^{(\rho_1,\rho_2)} \Theta\|_{L_x^2} + \left\|\frac{1}{r^2} A_\theta^{(\rho_1)} A_\theta^{(\rho_2)} \Theta\right\|_{L_x^2} \lesssim \|\psi_1\|_{\dot{H}_x^{1/2}} \|\psi_2\|_{\dot{H}_x^{1/2}} \|\Theta\|_{\dot{H}_x^{1/2}} \min_{\tau \in S_2} \|\psi_{\tau(1)}\|_{\dot{H}_x^{1/2}} \|\psi_{\tau(2)}\|_{L_x^2}. \tag{68}$$

Proof. We start with

$$\|A_0^{(\rho_1,\rho_2)} \Theta\|_{L_x^2} \lesssim \|r^{1/2} A_0^{(\rho_1,\rho_2)}\|_{L_x^\infty} \|r^{-1/2} \Theta\|_{L_x^2} \lesssim \|r^{1/2} A_0^{(\rho_1,\rho_2)}\|_{L_x^\infty} \|\Theta\|_{\dot{H}_x^{1/2}}$$

and then appeal to (67) with $p = q = 2$.

Similarly,

$$\left\|\frac{1}{r^2} A_\theta^{(\rho_1)} A_\theta^{(\rho_2)} \Theta\right\|_{L_x^2} \lesssim \|r^{-1} A_\theta^{(\rho_1)}\|_{L_x^\infty} \|r^{-1/2} A_\theta^{(\rho_2)}\|_{L_x^\infty} \|r^{-1/2} \Theta\|_{L_x^2}$$

$$\lesssim \|\psi_1\|_{\dot{H}_x^{1/2}}^2 \|\psi_2\|_{\dot{H}_x^{1/2}} \|\psi_2\|_{L_x^2} \|\Theta\|_{\dot{H}_x^{1/2}},$$

where we have used (66) and (65) with $p = q = 2$ and Hardy's inequality. Finally, we may repeat the estimate but with the roles of ψ_1 and ψ_2 reversed. $\qquad\square$

Now we introduce (see (28) to compare)

$$a_{\rho_1,\rho_2}(t,r) := A_0^{(\rho_1,\rho_2)}(t,r) + \tfrac{1}{r^2} A_\theta^{(\rho_1)}(t,r) A_\theta^{(\rho_2)}(t,r). \tag{69}$$

For the definitions of the terms on the right-hand side, see the equations and comments from (57) to (58).

Corollary 4.9. *Suppose $\rho_j = |\psi_j|^2$ for $j = 1, 2$. Then*

$$\|a_{\rho_1,\rho_2}\psi_3\|_{L_x^2} \lesssim \|\psi_1\|_{\dot{H}_x^{1/2}} \|\psi_2\|_{\dot{H}_x^{1/2}} \min_{\tau \in S_3} \|\psi_{\tau(1)}\|_{\dot{H}_x^{1/2}} \|\psi_{\tau(2)}\|_{\dot{H}_x^{1/2}} \|\psi_{\tau(3)}\|_{L_x^2}, \tag{70}$$

where S_3 denotes the set of permutations on three elements.

Proof. For all but two permutations, the estimate follows from (68). To establish the estimate for the remaining two cases, we need L_x^∞ bounds on $A_0^{(\rho_1,\rho_2)}$ and $\frac{1}{r^2} A_\theta^{(\rho_1)} A_\theta^{(\rho_2)}$. Using the second estimate of (62) twice and Sobolev embedding, we obtain

$$\left\| \tfrac{1}{r^2} A_\theta^{(\rho_1)} A_\theta^{(\rho_2)} \right\|_{L_x^\infty} \leq \left\| \tfrac{1}{r} A_\theta^{(\rho_1)} \right\|_{L_x^\infty} \left\| \tfrac{1}{r} A_\theta^{(\rho_2)} \right\|_{L_x^\infty} \lesssim \|\psi_1\|_{L_x^4}^2 \|\psi_2\|_{L_x^4}^2 \lesssim \|\psi_1\|_{\dot{H}_x^{1/2}}^2 \|\psi_2\|_{\dot{H}_x^{1/2}}^2.$$

To bound $A_0^{(\rho_1,\rho_2)}$, we proceed in a manner similar to that of the second estimate of (63) and (67). In particular, invoking (66) with $q = 2$ and Hardy, we obtain

$$\|A_0^{(\rho_1,\rho_2)}\|_{L_x^\infty} = \left\| \int_r^\infty s^{-1} A_\theta^{(\rho_1)} s^{-1} |\psi_2|^2 s \, ds \right\|_{L_x^\infty} \lesssim \|r^{-1} A_\theta^{(\rho_1)}\|_{L_x^\infty} \|r^{-1/2}\psi_2\|_{L_x^\infty}^2 \lesssim \|\psi_1\|_{\dot{H}_x^{1/2}}^2 \|\psi_2\|_{\dot{H}_x^{1/2}}^2. \qquad \square$$

Remark 4.10. From the proofs of these estimates, we see that the limiting factor in lowering the regularity of the unconditional uniqueness result lies in the interaction part, which requires $s = \frac{2}{3}$ rather than the $s = \frac{1}{2}$ required for the nonlinear part. By using negative-regularity Sobolev spaces, [Hong, Taliaferro, and Xie 2014] lowers the regularity required for the interaction part. Such a procedure does not seem to work, at least directly, for the problem at hand. This is because one would need to obtain the same negative-order Sobolev index in the right-hand side of (70) for the purpose of moving the term arising from controlling the nonlinear part back over to the left-hand side (see the argument following the proof of Theorem 2.4).

Acknowledgment

References

[Adami, Golse, and Teta 2007] R. Adami, F. Golse, and A. Teta, "Rigorous derivation of the cubic NLS in dimension one", *J. Stat. Phys.* **127**:6 (2007), 1193–1220. MR 2008i:82055 Zbl 1118.81021

[Ammari and Nier 2008] Z. Ammari and F. Nier, "Mean field limit for bosons and infinite dimensional phase-space analysis", *Ann. Henri Poincaré* **9**:8 (2008), 1503–1574. MR 2009m:81118 Zbl 1171.81014

[Ammari and Nier 2011] Z. Ammari and F. Nier, "Mean field propagation of Wigner measures and BBGKY hierarchies for general bosonic states", *J. Math. Pures Appl.* (9) **95**:6 (2011), 585–626. MR 2012e:81123 Zbl 1251.81062

[Arkeryd, Caprino, and Ianiro 1991] L. Arkeryd, S. Caprino, and N. Ianiro, "The homogeneous Boltzmann hierarchy and statistical solutions to the homogeneous Boltzmann equation", *J. Stat. Phys.* **63**:1-2 (1991), 345–361. MR 92m:82108

[Beckner 2014] W. Beckner, "Multilinear embedding-convolution estimates on smooth submanifolds", *Proc. Amer. Math. Soc.* **142**:4 (2014), 1217–1228. MR 3162244 Zbl 06269482

[Bejenaru, Ionescu, Kenig, and Tataru 2013] I. Bejenaru, A. Ionescu, C. E. Kenig, and D. Tataru, "Equivariant Schrödinger maps in two spatial dimensions", *Duke Math. J.* **162**:11 (2013), 1967–2025. MR 3090782 Zbl 06206217

[Benedetto, Castella, Esposito, and Pulvirenti 2005] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti, "On the weak-coupling limit for bosons and fermions", *Math. Models Methods Appl. Sci.* **15**:12 (2005), 1811–1843. MR 2007a:82035 Zbl 1154.82314

[Benedetto, Castella, Esposito, and Pulvirenti 2006] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti, "Some considerations on the derivation of the nonlinear quantum Boltzmann equation, II: The low density regime", *J. Stat. Phys.* **124**:2-4 (2006), 951–996. MR 2008g:82106 Zbl 1134.82029

[Benedetto, Castella, Esposito, and Pulvirenti 2008] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti, "From the *N*-body Schrödinger equation to the quantum Boltzmann equation: a term-by-term convergence result in the weak coupling regime", *Comm. Math. Phys.* **277**:1 (2008), 1–44. MR 2008i:82089 Zbl 1148.82022

[Benedikter, Oliveira, and Schlein 2012] N. Benedikter, G. de Oliveira, and B. Schlein, "Quantitative derivation of the Gross–Pitaevskii equation", preprint, 2012. To appear in *Comm. Pure Appl. Math.* arXiv 1208.0373

[Bergé, De Bouard, and Saut 1995] L. Bergé, A. De Bouard, and J.-C. Saut, "Blowing up time-dependent solutions of the planar, Chern–Simons gauged nonlinear Schrödinger equation", *Nonlinearity* **8**:2 (1995), 235–253. MR 96b:81025 Zbl 0822.35125

[Byeon, Huh, and Seok 2012] J. Byeon, H. Huh, and J. Seok, "Standing waves of nonlinear Schrödinger equations with the gauge field", *J. Funct. Anal.* **263**:6 (2012), 1575–1608. MR 2948224 Zbl 1248.35193

[Cercignani, Illner, and Pulvirenti 1994] C. Cercignani, R. Illner, and M. Pulvirenti, *The mathematical theory of dilute gases*, Applied Mathematical Sciences **106**, Springer, New York, 1994. MR 96g:82046 Zbl 0813.76001

[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

[Deser, Jackiw, and Templeton 1982] S. Deser, R. Jackiw, and S. Templeton, "Topologically massive gauge theories", *Ann. Physics* **140**:2 (1982), 372–411. MR 84j:81128

[Dunford and Schwartz 1988] N. Dunford and J. T. Schwartz, *Linear operators, I: General theory*, Wiley, New York, 1988. MR 90g:47001a Zbl 0635.47001

[Elgart, Erdős, Schlein, and Yau 2006] A. Elgart, L. Erdős, B. Schlein, and H.-T. Yau, "Gross–Pitaevskii equation as the mean field limit of weakly coupled bosons", *Arch. Ration. Mech. Anal.* **179**:2 (2006), 265–283. MR 2007b:81310 Zbl 1086.81035

[Erdős and Yau 2001] L. Erdős and H.-T. Yau, "Derivation of the nonlinear Schrödinger equation from a many body Coulomb system", *Adv. Theor. Math. Phys.* **5**:6 (2001), 1169–1205. MR 2004c:82075 Zbl 1014.81063

[Erdős, Salmhofer, and Yau 2004] L. Erdős, M. Salmhofer, and H.-T. Yau, "On the quantum Boltzmann equation", *J. Stat. Phys.* **116**:1-4 (2004), 367–380. MR 2006b:82086 Zbl 1142.82346

[Erdős, Schlein, and Yau 2006] L. Erdős, B. Schlein, and H.-T. Yau, "Derivation of the Gross–Pitaevskii hierarchy for the dynamics of Bose-Einstein condensate", *Comm. Pure Appl. Math.* **59**:12 (2006), 1659–1741. MR 2007k:82070 Zbl 1122.82018

[Erdős, Schlein, and Yau 2007] L. Erdős, B. Schlein, and H.-T. Yau, "Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems", *Invent. Math.* **167**:3 (2007), 515–614. MR 2007m:81258

[Erdős, Schlein, and Yau 2009] L. Erdős, B. Schlein, and H.-T. Yau, "Rigorous derivation of the Gross–Pitaevskii equation with a large interaction potential", *J. Amer. Math. Soc.* **22**:4 (2009), 1099–1156. MR 2010g:82041 Zbl 1207.82031

[Erdős, Schlein, and Yau 2010] L. Erdős, B. Schlein, and H.-T. Yau, "Derivation of the Gross–Pitaevskii equation for the dynamics of Bose-Einstein condensate", *Ann. of Math.* (2) **172**:1 (2010), 291–370. MR 2011g:82068 Zbl 1204.82028

[Ezawa, Hotta, and Iwazaki 1991a] Z. F. Ezawa, M. Hotta, and A. Iwazaki, "Breathing vortex solitons in nonrelativistic Chern–Simons gauge theory", *Phys. Rev. Lett.* **67**:4 (1991), 411–414. MR 92f:81100 Zbl 0990.81560

[Ezawa, Hotta, and Iwazaki 1991b] Z. F. Ezawa, M. Hotta, and A. Iwazaki, "Nonrelativistic Chern–Simons vortex solitons in external magnetic field", *Phys. Rev. D* (3) **44**:2 (1991), 452–463.

[Fröhlich, Knowles, and Schwarz 2009] J. Fröhlich, A. Knowles, and S. Schwarz, "On the mean-field limit of bosons with Coulomb two-body interaction", *Comm. Math. Phys.* **288**:3 (2009), 1023–1059. MR 2010b:82039 Zbl 1177.82016

[Gallagher, Saint-Raymond, and Texier 2013] I. Gallagher, L. Saint-Raymond, and B. Texier, *From Newton to Boltzmann: hard spheres and short-range potentials*, European Mathematical Society, Zürich, 2013. MR 3157048 Zbl 06248909

[Gressman, Sohinger, and Staffilani 2014] P. Gressman, V. Sohinger, and G. Staffilani, "On the uniqueness of solutions to the periodic 3D Gross–Pitaevskii hierarchy", *J. Funct. Anal.* **266**:7 (2014), 4705–4764. MR 3170216 Zbl 06320691

[Grillakis and Machedon 2013] M. G. Grillakis and M. Machedon, "Pair excitations and the mean field approximation of interacting bosons, I", *Comm. Math. Phys.* **324**:2 (2013), 601–636. MR 3117522 Zbl 1277.82034

[Grillakis and Margetis 2008] M. G. Grillakis and D. Margetis, "A priori estimates for many-body Hamiltonian evolution of interacting boson system", *J. Hyperbolic Differ. Equ.* **5**:4 (2008), 857–883. MR 2010f:35321 Zbl 1160.35357

[Grillakis, Machedon, and Margetis 2010] M. G. Grillakis, M. Machedon, and D. Margetis, "Second-order corrections to mean field evolution of weakly interacting bosons. I", *Comm. Math. Phys.* **294**:1 (2010), 273–301. MR 2011j:81374 Zbl 1208.82030

[Grillakis, Machedon, and Margetis 2011] M. G. Grillakis, M. Machedon, and D. Margetis, "Second-order corrections to mean field evolution of weakly interacting bosons. II", *Adv. Math.* **228**:3 (2011), 1788–1815. MR 2012m:81215 Zbl 1226.82033

[Hong, Taliaferro, and Xie 2014] Y. Hong, K. Taliaferro, and Z. Xie, "Unconditional uniqueness of the cubic Gross–Pitaevskii hierarchy with low regularity", preprint, 2014. arXiv 1402.5347

[Hudson and Moody 1976] R. L. Hudson and G. R. Moody, "Locally normal symmetric states and an analogue of de Finetti's theorem", *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **33**:4 (1976), 343–351. MR 53 #1280 Zbl 0304.60001

[Huh 2013] H. Huh, "Energy solution to the Chern–Simons–Schrödinger equations", *Abstr. Appl. Anal.* **2013** (2013), Art. ID #590653. MR 3035224 Zbl 1276.35138

[Jackiw and Pi 1990] R. Jackiw and S.-Y. Pi, "Classical and quantal nonrelativistic Chern–Simons theory", *Phys. Rev. D* (3) **42**:10 (1990), 3500–3513. MR 1084552

[Jackiw and Pi 1991] R. Jackiw and S.-Y. Pi, "Time-dependent Chern–Simons solitons and their quantization", *Phys. Rev. D* (3) **44**:8 (1991), 2524–2532. MR 93d:81108

[Jackiw and Pi 1992] R. Jackiw and S.-Y. Pi, "Self-dual Chern–Simons solitons", *Progr. Theoret. Phys. Suppl.* **107** (1992), 1–40. MR 94d:81114

[Jackiw and Templeton 1981] R. Jackiw and S. Templeton, "How super-renormalizable interactions cure their infrared divergences", *Phys. Rev. D* (3) **23**:10 (1981), 2291–2304.

[Jackiw and Weinberg 1990] R. Jackiw and E. J. Weinberg, "Self-dual Chern–Simons vortices", *Phys. Rev. Lett.* **64**:19 (1990), 2234–2237. MR 91a:81117 Zbl 1050.81595

[Jackiw, Pi, and Weinberg 1991] R. Jackiw, S.-Y. Pi, and E. J. Weinberg, "Topological and nontopological solitons in relativistic and nonrelativistic Chern–Simons theory", pp. 573–588 in *Particles, strings and cosmology* (Boston, 1990), edited by P. Nath and S. Reucroft, World Scientific, River Edge, NJ, 1991. MR 1170608

[King 1975] F. G. King, *BBGKY hierarchy for positive potentials*, Ph.D. thesis, University of California, Berkeley, CA, 1975, Available at http://search.proquest.com/docview/302766232. MR 2625983

[Kirkpatrick, Schlein, and Staffilani 2011] K. Kirkpatrick, B. Schlein, and G. Staffilani, "Derivation of the two-dimensional nonlinear Schrödinger equation from many body quantum dynamics", *Amer. J. Math.* **133**:1 (2011), 91–130. MR 2012f:81079 Zbl 1208.81080

[Klainerman and Machedon 2008] S. Klainerman and M. Machedon, "On the uniqueness of solutions to the Gross–Pitaevskii hierarchy", *Comm. Math. Phys.* **279**:1 (2008), 169–185. MR 2009a:35236 Zbl 1147.37034

[Knowles and Pickl 2010] A. Knowles and P. Pickl, "Mean-field dynamics: singular potentials and rate of convergence", *Comm. Math. Phys.* **298**:1 (2010), 101–138. MR 2011j:81096 Zbl 1213.81180

[L. Chen, Lee, and Schlein 2011] L. Chen, J. O. Lee, and B. Schlein, "Rate of convergence towards Hartree dynamics", *J. Stat. Phys.* **144**:4 (2011), 872–903. MR 2012h:81078 Zbl 1227.82046

[Lanford 1975] O. E. Lanford, III, "Time evolution of large classical systems", pp. 1–111 in *Dynamical systems, theory and applications* (Seattle, WA, 1974), edited by J. Moser, Lecture Notes in Phys. 38, Springer, Berlin, 1975. MR 57 #18653 Zbl 0329.70011

[Lewin, Nam, and Rougerie 2014] M. Lewin, P. T. Nam, and N. Rougerie, "Derivation of Hartree's theory for generic mean-field Bose systems", *Adv. Math.* **254** (2014), 570–621. MR 3161107 Zbl 06285011

[Liu and Smith 2014] B. Liu and P. Smith, "Global wellposedness of the equivariant Chern–Simons–Schrödinger equation", preprint, 2014. arXiv 1312.5567

[Liu, Smith, and Tataru 2012] B. Liu, P. Smith, and D. Tataru, "Local wellposedness of Chern–Simons–Schrödinger", preprint, 2012. To appear in *Int. Math. Res. Not.* arXiv 1212.1476

[Martina, Pashaev, and Soliani 1993] L. Martina, O. K. Pashaev, and G. Soliani, "Self-dual Chern–Simons solitons in nonlinear σ-models", *Modern Phys. Lett. A* **8**:34 (1993), 3241–3250. MR 94g:81138 Zbl 1021.81661

[Michelangeli and Schlein 2012] A. Michelangeli and B. Schlein, "Dynamical collapse of boson stars", *Comm. Math. Phys.* **311**:3 (2012), 645–687. MR 2909759 Zbl 1242.85007

[Paul and Khare 1986] S. K. Paul and A. Khare, "Charged vortices in an abelian Higgs model with Chern–Simons term", *Phys. Lett. B* **174**:4 (1986), 420–422. MR 87j:81269a

[Pickl 2011] P. Pickl, "A simple derivation of mean field limits for quantum systems", *Lett. Math. Phys.* **97**:2 (2011), 151–164. MR 2012i:81299 Zbl 1242.81150

[R. M. Chen and Spirn 2009] R. M. Chen and D. Spirn, "Symmetric Chern–Simons–Higgs vortices", *Comm. Math. Phys.* **285**:3 (2009), 1005–1031. MR 2010e:58021 Zbl 1171.81020

[Rodnianski and Schlein 2009] I. Rodnianski and B. Schlein, "Quantum fluctuations and rate of convergence towards mean field dynamics", *Comm. Math. Phys.* **291**:1 (2009), 31–61. MR 2011d:82059 Zbl 1186.82051

[Sohinger 2014a] V. Sohinger, "Local existence of solutions to randomized Gross–Pitaevskii hierarchies", preprint, 2014. To appear in *Trans. Amer. Math. Soc.* arXiv 1401.0326

[Sohinger 2014b] V. Sohinger, "A rigorous derivation of the defocusing cubic nonlinear Schrödinger equation on $\mathbb{T}^3$ from the dynamics of many-body quantum systems", *Ann. Inst. H. Poincaré Anal. Non Linéaire* (online publication October 2014).

[Sohinger and Staffilani 2014] V. Sohinger and G. Staffilani, "Randomization and the Gross–Pitaevskii hierarchy", preprint, 2014. arXiv 1308.3714

[Spohn 1980] H. Spohn, "Kinetic equations from Hamiltonian dynamics: Markovian limits", *Rev. Modern Phys.* **52**:3 (1980), 569–615. MR 81e:82010

[Størmer 1969] E. Størmer, "Symmetric states of infinite tensor products of C^*-algebras", *J. Funct. Anal.* **3** (1969), 48–68. MR 39 #3327 Zbl 0167.43403

[T. Chen and Pavlović 2010] T. Chen and N. Pavlović, "On the Cauchy problem for focusing and defocusing Gross–Pitaevskii hierarchies", *Discrete Contin. Dyn. Syst.* **27**:2 (2010), 715–739. MR 2011f:35317 Zbl 1190.35207

[T. Chen and Pavlović 2011] T. Chen and N. Pavlović, "The quintic NLS as the mean field limit of a boson gas with three-body interactions", *J. Funct. Anal.* **260**:4 (2011), 959–997. MR 2012b:82066 Zbl 1213.35368

[T. Chen and Pavlović 2014] T. Chen and N. Pavlović, "Derivation of the cubic NLS and Gross–Pitaevskii hierarchy from manybody dynamics in $d = 3$ based on spacetime norms", *Ann. Henri Poincaré* **15**:3 (2014), 543–588. MR 3165917 Zbl 06290593

[T. Chen, Hainzl, Pavlović, and Seiringer 2014] T. Chen, C. Hainzl, N. Pavlović, and R. Seiringer, "Unconditional uniqueness for the cubic Gross–Pitaevskii hierarchy via quantum de Finetti", preprint, 2014. To appear in *Comm. Pure Appl. Math.* arXiv 1307.3168

[T. Chen, Pavlović, and Tzirakis 2012] T. Chen, N. Pavlović, and N. Tzirakis, "Multilinear Morawetz identities for the Gross–Pitaevskii hierarchy", pp. 39–62 in *Recent advances in harmonic analysis and partial differential equations* (Statesboro, GA, 2011 and Baltimore, MD, 2011), edited by A. R. Nahmod et al., Contemp. Math. **581**, American Mathematical Society, Providence, RI, 2012. MR 3013052

[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

[de Vega and Schaposnik 1986a] H. J. de Vega and F. A. Schaposnik, "Electrically charged vortices in nonabelian gauge theories with Chern–Simons term", *Phys. Rev. Lett.* **56**:24 (1986), 2564–2566. MR 87f:81099

[de Vega and Schaposnik 1986b] H. J. de Vega and F. A. Schaposnik, "Vortices and electrically charged vortices in nonabelian gauge theories", *Phys. Rev. D* (3) **34**:10 (1986), 3206–3213. MR 88a:81139

[Wilczek 1990] F. Wilczek, *Fractional statistics and anyon superconductivity*, World Scientific, Teaneck, NJ, 1990. MR 92c: 81001

[X. Chen 2011] X. Chen, "Classical proofs of Kato type smoothing estimates for the Schrödinger equation with quadratic potential in $\mathbb{R}^{n+1}$ with application", *Differential Integral Equations* **24**:3-4 (2011), 209–230. MR 2011m:35303 Zbl 1240.35429

[X. Chen 2012a] X. Chen, "Collapsing estimates and the rigorous derivation of the 2d cubic nonlinear Schrödinger equation with anisotropic switchable quadratic traps", *J. Math. Pures Appl.* (9) **98**:4 (2012), 450–478. MR 2968164 Zbl 1251.35144

[X. Chen 2012b] X. Chen, "Second order corrections to mean field evolution for weakly interacting bosons in the case of three-body interactions", *Arch. Ration. Mech. Anal.* **203**:2 (2012), 455–497. MR 2885567 Zbl 1256.35099

[X. Chen 2013] X. Chen, "On the rigorous derivation of the 3D cubic nonlinear Schrödinger equation with a quadratic trap", *Arch. Ration. Mech. Anal.* **210**:2 (2013), 365–408. MR 3101788 Zbl 1294.35132

[X. Chen and Holmer 2013a] X. Chen and J. Holmer, "Focusing quantum many-body dynamics: the rigorous derivation of the 1D focusing cubic nonlinear Schrödinger equation", preprint, 2013. arXiv 1308.3895

[X. Chen and Holmer 2013b] X. Chen and J. Holmer, "On the Klainerman–Machedon conjecture of the quantum BBGKY hierarchy with self-interaction", preprint, 2013. To appear in *J. Eur. Math. Soc.* arXiv 1303.5385

[X. Chen and Holmer 2013c] X. Chen and J. Holmer, "On the rigorous derivation of the 2D cubic nonlinear Schrödinger equation from 3D quantum many-body dynamics", *Arch. Ration. Mech. Anal.* **210**:3 (2013), 909–954. MR 3116008 Zbl 1288.35429

XUWEN CHEN: chenxuwen@math.brown.edu
Department of Mathematics, Brown University, 151 Thayer Street, Box 1917, Providence, RI 02912, United States

PAUL SMITH: smith@math.berkeley.edu
Mathematics Department, University of California, Berkeley, Evans Hall, Berkeley, CA 94720-3840, United States
Current address: *Google, 1600 Amphitheatre Parkway, Mountain View, CA 94043, United States*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.