

# SPECTRAL GAPS, ADDITIVE ENERGY, AND A FRACTAL UNCERTAINTY PRINCIPLE

SEMYON DYATLOV AND JOSHUA ZAHL

**ABSTRACT.** We obtain an essential spectral gap for  $n$ -dimensional convex co-compact hyperbolic manifolds with the dimension  $\delta$  of the limit set close to  $\frac{n-1}{2}$ . The size of the gap is expressed using the additive energy of stereographic projections of the limit set. This additive energy can in turn be estimated in terms of the constants in Ahlfors-David regularity of the limit set. Our proofs use new microlocal methods, in particular a notion of a fractal uncertainty principle.

In this paper we study essential spectral gaps for convex co-compact hyperbolic quotients  $M = \Gamma \backslash \mathbb{H}^n$ . To formulate our result in the simplest setting, consider  $n = 2$  and take the *Selberg zeta function* [Bo07, (10.1)]

$$Z_M(\lambda) = \prod_{\ell \in \mathcal{L}_M} \prod_{k=0}^{\infty} (1 - e^{-(s+k)\ell}), \quad s = \frac{1}{2} - i\lambda,$$

where  $\mathcal{L}_M$  consists of all lengths of primitive closed geodesics on  $M$  (with multiplicity). The spectral representation of  $Z_M$  implies that it has only finitely many singularities (that is, zeros and poles) in  $\{\operatorname{Im} \lambda > 0\}$ . The work of Patterson [Pa76b] and Sullivan [Su] shows that  $Z_M(\lambda)$  has no singularities in  $\{\operatorname{Im} \lambda > \delta - \frac{1}{2}\}$  and a simple zero at  $i(\delta - \frac{1}{2})$ , where  $\delta \in [0, 1]$  is the dimension of the limit set of the group (see (5.2)). Therefore  $Z_M$  has finitely many singularities in  $\{\operatorname{Im} \lambda > -\max(0, \frac{1}{2} - \delta)\}$ .

For  $\delta \in (0, \frac{1}{2}]$ , Naud [Na] obtained the stronger statement that  $Z_M$  has finitely many singularities in  $\{\operatorname{Im} \lambda > -\beta\}$  for some  $\beta$  strictly greater than  $\frac{1}{2} - \delta$ . Naud's result, generalized to higher dimensional quotients by Stoyanov [St11], is based on the method of Dolgopyat [Do] and does not specify the size of the improvement. Our first result in particular gives explicit estimates on the value of  $\beta$  when  $\delta = \frac{1}{2}$ :

**Theorem 1.** *Let  $M = \Gamma \backslash \mathbb{H}^2$  be a convex co-compact hyperbolic surface. Then for each  $\varepsilon > 0$ , the function  $Z_M$  has finitely many singularities in  $\{\operatorname{Im} \lambda > -\beta + \varepsilon\}$ , where*

$$\beta = \frac{3}{8} \left( \frac{1}{2} - \delta \right) + \frac{\beta_E}{16}, \quad \beta_E := \delta \exp \left[ -\mathbf{K}(1 - \delta)^{-28} (1 + \log^{14} \mathbf{C}) \right]. \quad (1.1)$$

Here  $\mathbf{K} > 0$  is a global constant and  $\mathbf{C} \geq 1$  is the constant in the Ahlfors-David regularity for the limit set  $\Lambda_\Gamma$ , see (1.24).

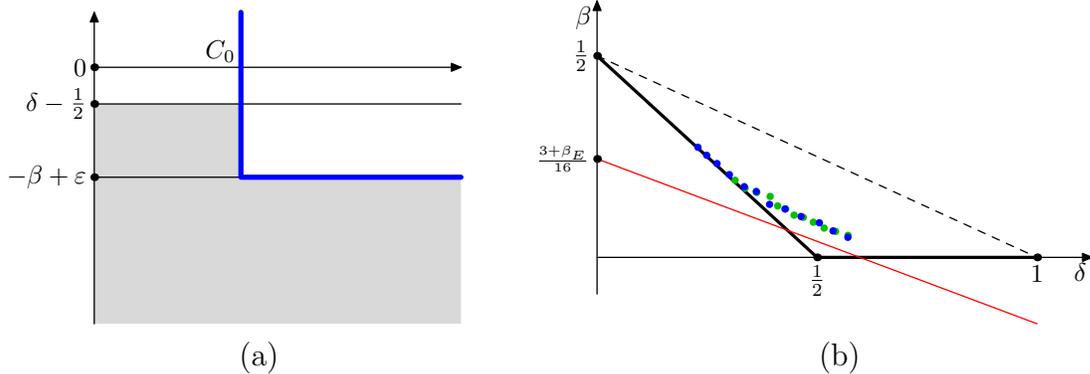


FIGURE 1. (a) Resonance free regions (in white) in dimension 2 for  $\delta < \frac{1}{2}$  given by the Patterson–Sullivan gap (1.2) and an essential spectral gap of Theorem 2 (outlined in blue). (b) Essential spectral gaps  $\beta$  in dimension 2 depending on  $\delta$ : the standard gap (in bold) and the Jakobson–Naud conjecture (dashed). The circles correspond to numerically computed spectral gaps for symmetric 3-funneled surfaces (blue) and 4-funneled surfaces (green) from [BoWe, Figure 14] (specifically,  $G_{100}^4$  in the notation of [BoWe]). The red line is the gap given by Theorem 1 for a *sample choice* of  $\beta_E := \frac{1}{2}$ .

See Figure 1. We say that  $M$  has an *essential spectral gap* of size  $\beta$ . Theorem 2 improves over the standard gap  $\beta_{\text{std}} = \max(0, \frac{1}{2} - \delta)$  for all surfaces with  $\delta = \frac{1}{2}$ . In fact, we show on the example of three-funnel surfaces that the regularity constant  $\mathbf{C}$  is bounded when the surface  $M$  varies in a compact set in the moduli space, and thus (1.1) improves over  $\beta_{\text{std}}$  for surfaces close to those with  $\delta = \frac{1}{2}$  – see Theorem 7 in §7.3. This includes some surfaces with  $\delta > \frac{1}{2}$ , which to our knowledge provide the first examples of fractal chaotic systems where the pressure  $P(\frac{1}{2})$  is positive but there is an essential spectral gap below the real line. On the other hand our methods cannot improve over  $\beta_{\text{std}}$  even in the most favorable situation unless  $\delta \in (\frac{5}{11}, \frac{3}{5})$  – see the remark following Theorem 4.

A numerical investigation of the gap was done by Borthwick [Bo14, §7] and Borthwick–Weich [BoWe] (see Figure 1(b)). Both [BoWe, Figure 14] and the experimental results of [BWPSKZ, Figure 4] suggest that the improvement  $\beta - \beta_{\text{std}}$  should indeed be larger when  $\delta \approx \frac{1}{2}$  than for other values of  $\delta$ .

In fact, our results apply to convex co-compact quotients  $M$  of any dimension  $n$  and give bounds on the *scattering resolvent*

$$R(\lambda) = \left( -\Delta - \frac{(n-1)^2}{4} - \lambda^2 \right)^{-1} : L_{\text{comp}}^2(M) \rightarrow H_{\text{loc}}^2(M), \quad \lambda \in \mathbb{C},$$

which is the meromorphic continuation of the  $L^2$  resolvent from the upper half-plane – see §4.2. The standard Patterson–Sullivan gap in this setting is [Bo07, §14.4]

$$R(\lambda) \text{ has only finitely many poles in } \left\{ \operatorname{Im} \lambda \geq -\max\left(0, \frac{n-1}{2} - \delta\right) \right\}. \quad (1.2)$$

The poles of  $R(\lambda)$ , called *resonances*, are related to the scattering poles and to the zeroes of  $Z_M$  as proved by Bunke–Olbrich [BuOl99] and Patterson–Perry [PaPe]; see Borthwick [Bo07, Chapter 10] for an expository proof and the history of the subject. Therefore Theorem 1 is a direct corollary of the following stronger statement:

**Theorem 2.** *Under the assumptions of Theorem 1, we have the resolvent estimate*

$$\|\chi R(\lambda)\chi\|_{L^2 \rightarrow L^2} \leq C|\lambda|^{-1-2\min(0, \operatorname{Im} \lambda)+\varepsilon}, \quad |\lambda| > C_0, \operatorname{Im} \lambda \in [-\beta + \varepsilon, 1], \quad (1.3)$$

where  $\beta$  is given by (1.1),  $\varepsilon > 0$  and  $\chi \in C_0^\infty(M)$  are arbitrary, the constant  $C_0$  depends on  $\varepsilon$ , and the constant  $C$  depends on  $\varepsilon, \chi$ .

Spectral gaps for the special case of arithmetic quotients have recently found important applications to diophantine problems, see Bourgain–Gamburd–Sarnak [BGS] and Magee–Oh–Winter [MOW]. For these applications one needs a uniform resonance free region for congruence subgroups of  $\Gamma$ ; a uniform logarithmic region was obtained in [BGS] and a uniform gap by Oh–Winter [OhWi].

Our results have the following features compared to previous works on spectral gaps:

- The size of the gap is more explicit, expressed in terms of additive energy of the limit set (Theorem 4) or the constants in Ahlfors–David regularity of this set. Compared to Dolgopyat’s method, we decouple analytical aspects of the problem from combinatorial ones. This makes it more feasible to compute the size of the gap for specific hyperbolic manifolds.
- We obtain a polynomial resolvent bound (1.3), rather than just a resonance free strip. This can be used to obtain polynomial bounds on other objects such as Eisenstein functions.
- We rely on  $C^\infty$  microlocal analysis (for instance, nowhere using explicitly that  $M$  is analytic); this gives hope that our strategy may apply to more general classes of manifolds.

Regarding the last item on the above list, we make the following conjecture which would improve over the *pressure gap* studied in more general cases by Ikawa [Ik], Gaspard–Rice [GaRi], and Nonnenmacher–Zworski [NoZw09].

**Conjecture.** *Let  $(M, g)$  be a convex co-compact hyperbolic surface with  $\delta = \frac{1}{2}$ . Then all sufficiently small smooth compactly supported metric perturbations of  $(M, g)$  satisfy (1.3) for some  $\beta > 0$ .*

This conjecture is related to the improved gaps obtained for scattering by several convex obstacles by Petkov–Stoyanov [PeSt] and Stoyanov [St12] using the methods of [Do].

We now describe the scheme of proof of Theorem 2:

**1.1. Spectral gaps via a fractal uncertainty principle.** We first reduce the estimate (1.3) to a fractal uncertainty principle. To state it, let  $\Lambda_\Gamma \subset \mathbb{S}^{n-1}$  be the limit set of the group  $\Gamma$ ,  $M = \Gamma \backslash \mathbb{H}^n$  (see (4.11)) and denote by  $\mathbb{1}_{\Lambda_\Gamma(\alpha)}$  the indicator function of the set

$$\Lambda_\Gamma(\alpha) = \{y \in \mathbb{S}^{n-1} \mid d(y, \Lambda_\Gamma) \leq \alpha\} \quad (1.4)$$

where  $d(y, y') = |y - y'|$  denotes the Euclidean distance function on  $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ . Note that the Minkowski dimension of  $\Lambda_\Gamma$  is equal to  $\delta$ , therefore (see (5.3))

$$\alpha^{n-1-\delta}/C \leq \mu_L(\Lambda_\Gamma(\alpha)) \leq C\alpha^{n-1-\delta}, \quad \alpha \in (0, 1) \quad (1.5)$$

where  $\mu_L$  denotes the Lebesgue measure on  $\mathbb{S}^{n-1}$ .

Define the operator  $\mathcal{B}_\chi = \mathcal{B}_\chi(h) : L^2(\mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{S}^{n-1})$  by

$$\mathcal{B}_\chi v(y) = (2\pi h)^{\frac{1-n}{2}} \int_{\mathbb{S}^{n-1}} |y - y'|^{2i/h} \chi(y, y') v(y') dy' \quad (1.6)$$

where  $dy'$  is the standard volume form on  $\mathbb{S}^{n-1}$  and  $\chi \in C_0^\infty(\mathbb{S}_\Delta^{n-1})$ , where

$$\mathbb{S}_\Delta^{n-1} = \{(y, y') \in \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \mid y \neq y'\}. \quad (1.7)$$

**Definition 1.1.** *We say that  $\Lambda_\Gamma$  satisfies the **fractal uncertainty principle** with exponent  $\beta > 0$ , if for each  $\varepsilon > 0$  there exists  $\rho \in (0, 1)$  such that*

$$\| \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \mathcal{B}_\chi(h) \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \|_{L^2(\mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{S}^{n-1})} \leq C h^{\beta-\varepsilon}, \quad h \in (0, 1) \quad (1.8)$$

for every  $h$ -independent constant  $C_1$  and function  $\chi \in C_0^\infty(\mathbb{S}_\Delta^{n-1})$ , and some  $C$  depending on  $C_1, \chi$ .

**Remark.** The fractal uncertainty principle always holds with  $\beta = \max(0, \frac{n-1}{2} - \delta)$ , see (5.1) and (5.4). On the other hand, by (5.5) the maximal  $\beta$  for which (1.8) can be true is  $\beta = \frac{n-1}{2} - \frac{\delta}{2}$ , which (in dimension 2) is exactly the value of the essential spectral gap conjectured by Jakobson–Naud [JaNa].

To explain how the estimate (1.8) represents an uncertainty principle associated to the set  $\Lambda_\Gamma$ , we consider the extremal case  $\rho = 1$ , put  $C_1 = 1$ , and cover  $\Lambda_\Gamma(h)$  by a collection of balls of radius  $h$  centered at some points  $y_1, \dots, y_N \in \Lambda_\Gamma$ , where  $N \sim h^{-\delta}$  by (1.5). Then for each  $v \in L^2(\mathbb{S}^{n-1})$ , the function  $\mathcal{B}_\chi(h) \mathbb{1}_{\Lambda_\Gamma(h)} v$  microlocally concentrates (see (1.12) below) in an  $h$ -neighborhood of the union of ‘horizontal’ Lagrangian leaves

$$\bigcup_{j=1}^N \{(y, \partial_y \log(|y - y_j|^2)) \mid (y, y_j) \in \text{supp } \chi\} \subset T^*\mathbb{S}^{n-1}, \quad (1.9)$$

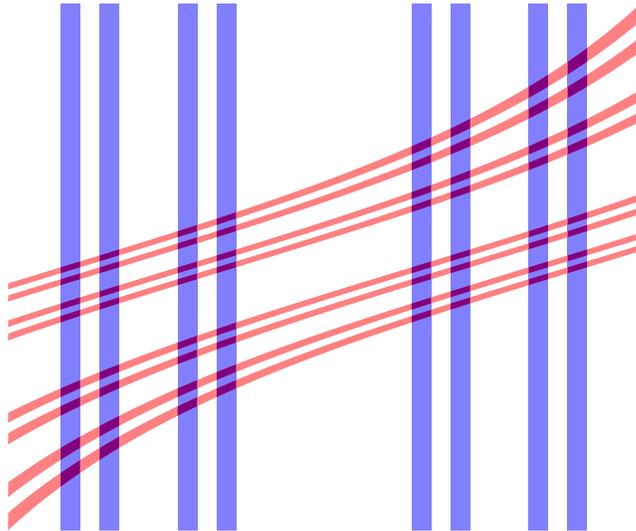


FIGURE 2. The horizontal leaves (1.9), in red, and the vertical leaves (1.10), in blue, for  $n = 2$ . The horizontal variable is  $y \in \mathbb{S}^1$  and the vertical variable is  $\eta$ . For the fractal uncertainty principle, the width of the distorted rectangles is slightly larger than  $h$ .

while the operator  $\mathbb{1}_{\Lambda_\Gamma(h)}$  microlocalizes to an  $h$ -neighborhood of the union of ‘vertical’ Lagrangian leaves

$$\bigcup_{j=1}^N \{(y_j, \eta) \mid \eta \in T_{y_j}^* \mathbb{S}^{n-1}\} \subset T^* \mathbb{S}^{n-1}. \quad (1.10)$$

The estimate (1.8) with  $\beta > 0$  then says that no function can be perfectly localized to  $h$ -neighborhoods of both (1.9) and (1.10) – see Figure 2. Note that  $h$ -neighborhoods here cannot be replaced by, say,  $h^{1/2}$  neighborhoods since Gaussians provide examples of functions that concentrate  $h^{1/2}$  close to any fixed leaf of (1.9) and to any fixed leaf of (1.10). A related statement in the context of normally hyperbolic trapping was proved by Nonnenmacher–Zworski [NoZw15, Lemma 5.12].

If  $\Lambda_\Gamma$  satisfies the fractal uncertainty principle, then an essential spectral gap is given by the following

**Theorem 3.** *Assume that  $\Lambda_\Gamma$  satisfies the fractal uncertainty principle with exponent  $\beta > 0$ . Then (1.3) holds, in particular  $R(\lambda)$  has finitely many poles in  $\{\text{Im } \lambda > -\beta + \varepsilon\}$  for each  $\varepsilon > 0$ .*

We outline the proof of the resonance free region of Theorem 3 (the resolvent bound follows directly from the argument). It suffices to show that for  $\text{Re } \lambda \gg 1$  and  $\text{Im } \lambda \geq -\beta + \varepsilon$ , there are no nontrivial *resonant states*, that is solutions to the equation

$$\left( -\Delta - \frac{(n-1)^2}{4} - \lambda^2 \right) u = 0 \quad (1.11)$$

which satisfy certain *outgoing* conditions asymptotically at the infinite ends of  $M$ .

Put  $h := (\operatorname{Re} \lambda)^{-1}$  and assume that  $u$  is  $L^2$ -normalized on a sufficiently large fixed compact subset of  $M$ . We study concentration of  $u$  in the *phase space*  $T^*M$  using semiclassical quantization

$$a \in C^\infty(T^*M) \mapsto \operatorname{Op}_h(a) : C^\infty(M) \rightarrow C^\infty(M) \quad (1.12)$$

where  $a$  satisfies certain growth conditions – see §2.

Let  $\Gamma_+ \subset T^*M \setminus 0$  be the *outgoing tail*, consisting of geodesics which are trapped backwards in time; define also the *incoming tail*  $\Gamma_- \subset T^*M \setminus 0$  (see (4.10)). The work of Vasy [Va13a, Va13b] near the infinite ends together with propagation of semiclassical singularities shows that  $u$  is microlocalized on  $\Gamma_+$  (see [BoMi, NoZw09] for related results in Euclidean scattering). More precisely, for an  $h$ -independent symbol  $a_+$ ,

$$\operatorname{supp}(1 - a_+) \cap \Gamma_+ = \emptyset \implies (1 - \operatorname{Op}_h(a_+))u = \mathcal{O}(h^\infty)_{C^\infty(M)}. \quad (1.13)$$

Moreover,  $u$  has positive mass near  $\Gamma_-$ ; more precisely, for  $h$ -independent  $a_-$

$$\operatorname{supp}(1 - a_-) \cap \Gamma_- = \emptyset \implies \|\operatorname{Op}_h(a_-)u\|_{L^2} \geq C^{-1} > 0. \quad (1.14)$$

(The statement (1.14) is not quite correct since  $\Gamma_-$  extends to the infinite ends of  $M$  and thus  $a_-$  cannot be compactly supported; however, we may argue in a fixed neighborhood of the trapped set  $K = \Gamma_+ \cap \Gamma_-$ . See Lemma 4.4 for precise statements.) The main idea of the proof is to replace  $h$ -independent symbols in (1.13) and (1.14) with symbols that concentrate  $h^\rho$  close to  $\Gamma_\pm$ :

$$d(\operatorname{supp}(1 - a_+), \Gamma_+) > h^\rho \implies (1 - \operatorname{Op}_h(a_+))u = \mathcal{O}(h^\infty)_{C^\infty(M)}, \quad (1.15)$$

$$d(\operatorname{supp}(1 - a_-), \Gamma_-) > h^\rho \implies \|\operatorname{Op}_h(a_-)u\|_{L^2} \geq C^{-1}h^{(-\operatorname{Im} \lambda)\rho}. \quad (1.16)$$

The constant  $\rho \in (0, 1)$  is taken very close to 1. See Lemma 4.6 for precise statements.

The proofs of (1.13) and (1.14) use propagation estimates for some  $h$ -independent time. The proofs of (1.15) and (1.16) use similar estimates for time  $t = \rho \log(1/h)$ , and the factor  $h^{(-\operatorname{Im} \lambda)\rho} = e^{(\operatorname{Im} \lambda)t}$  results from the imaginary part of the operator in (1.11).

However, the analysis for (1.15) and (1.16) is considerably more complicated since the symbols  $a_\pm$  have very rough behavior in the directions transversal to  $\Gamma_\pm$ , oscillating on the scale  $h^\rho$  – this corresponds to the fact that  $t$  is almost *twice* the Ehrenfest time (since the maximal expansion rate for the geodesic flow is equal to 1, the Ehrenfest time is just below  $\frac{1}{2} \log(1/h)$  – see for instance [DyGu14a, Proposition 3.9]). To solve this problem, we use the fact that  $\Gamma_+$  is foliated by the leaves of the weak unstable Lagrangian foliation  $L_u$ , while  $\Gamma_-$  is foliated by the leaves of the weak stable Lagrangian foliation  $L_s$ ; therefore, we can make  $a_+$  vary on scale 1 along  $L_u$  and  $a_-$  vary on the scale 1 along  $L_s$ . Then  $a_+$  and  $a_-$  can both be quantized to some operators  $\operatorname{Op}_h^{L_u}(a_+)$  and  $\operatorname{Op}_h^{L_s}(a_-)$ ; however, these operators will not be part of the same calculus – see §3 for details.

Next, the fractal uncertainty principle gives the following estimate for some  $a_{\pm}$  satisfying the conditions of (1.15), (1.16):

$$\|\mathrm{Op}_h(a_-)\mathrm{Op}_h(a_+)\|_{L^2 \rightarrow L^2} \leq Ch^{\beta-\varepsilon/2}. \quad (1.17)$$

To see this, we conjugate by a Fourier integral operator whose underlying canonical transformation maps an  $h^\rho$  neighborhood of  $\Gamma_+, \Gamma_-$  to an  $h^\rho$  neighborhood of (1.9), (1.10) respectively (strictly speaking, to the products of (1.9), (1.10) with  $(T^*\mathbb{R}^+)_w$ , where  $w$  corresponds to  $|\xi|_g$  and  $\partial_\theta$  corresponds to the generator of the geodesic flow). Under this conjugation,  $\mathrm{Op}_h(a_-)$  corresponds to  $\mathbb{1}_{\Lambda_\Gamma(h^\rho)}$  and  $\mathrm{Op}_h(a_+)$  corresponds to  $\mathcal{B}_\chi \mathbb{1}_{\Lambda_\Gamma(h^\rho)} \mathcal{B}_\chi^*$ , therefore (1.17) follows from (1.8). See §4.4 for details.

Gathering together (1.15), (1.16), and (1.17), and recalling that  $-\mathrm{Im} \lambda < \beta - \varepsilon$ , we obtain a contradiction for  $\rho$  close enough to 1 and  $h \ll 1$  (thus finishing the proof):

$$\begin{aligned} C^{-1}h^{(\beta-\varepsilon)\rho} &\leq C^{-1}h^{(-\mathrm{Im} \lambda)\rho} \leq \|\mathrm{Op}_h(a_-)u\|_{L^2} \\ &\leq \|\mathrm{Op}_h(a_-)\mathrm{Op}_h(a_+)u\|_{L^2} + \mathcal{O}(h^\infty) \leq Ch^{\beta-\varepsilon/2}. \end{aligned}$$

It would be interesting to see if Theorem 3 could be proved using transfer operator techniques such as the ones in [Na]. We however note that the microlocal argument presented above may be easier to adapt to a variable curvature situation (see the Conjecture above) and it also provides an explicit polynomial bound on the resolvent (1.3).

**1.2. Fractal uncertainty principle via additive energy.** As remarked before (following Definition 1.1), the fractal uncertainty principle holds with  $\beta = \frac{n-1}{2} - \delta$ . This corresponds to counting the total area of the intersections of  $h$ -neighborhoods of (1.9) and (1.10) (which in turn depend on  $\delta$  by (1.5)) and can be seen via an  $L^1 \rightarrow L^\infty$  norm bound on  $\mathcal{B}_\chi$ . On the other hand, an  $L^2 \rightarrow L^2$  norm bound on  $\mathcal{B}_\chi$  gives the fractal uncertainty principle with  $\beta = 0$ .

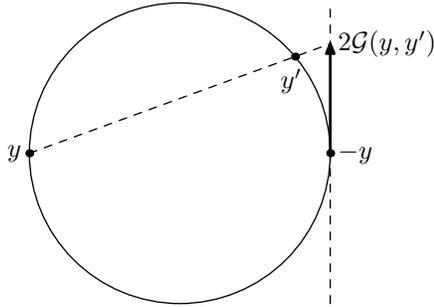
If we only use the volume bound (1.5), then no better value of  $\beta$  can be obtained – for a non-rigorous explanation, one may replace  $\Lambda_\Gamma(C_1h^\rho)$  in (1.8) by a ball of volume  $h^{n-1-\delta}$  in  $\mathbb{R}^{n-1}$ , replace  $\mathcal{B}_\chi$  by the semiclassical Fourier transform, and calculate the corresponding  $L^2 \rightarrow L^2$  norm.

To get a better exponent  $\beta$ , we thus have to use the fractal structure of  $\Lambda_\Gamma$ . More precisely, we will rely on the following combinatorial quantity:

**Definition 1.2.** For  $\mathcal{X} \subset \mathbb{R}^{n-1}$  and  $\alpha > 0$ , define the  $\alpha$ -**additive energy** of  $\mathcal{X}$  by

$$E_A(\mathcal{X}, \alpha) = \alpha^{4(1-n)} \mu_L(\{(\eta_1, \eta_2, \eta_3, \eta_4) \in \mathcal{X}(\alpha)^4 \mid |\eta_1 - \eta_2 + \eta_3 - \eta_4| \leq \alpha\})$$

where  $\mathcal{X}(\alpha)$  is the  $\alpha$ -neighborhood of  $\mathcal{X}$  and  $\mu_L$  is the Lebesgue measure. This definition trivially extends from  $\mathbb{R}^{n-1}$  to any  $n-1$  dimensional vector space with an inner product.

FIGURE 3. The stereographic projection map  $\mathcal{G}$ .

Additive energy is intimately connected with the additive structure of finite sets, and it is one of the central concepts in the field of additive combinatorics. See [TaVu06] for further information on additive energy and related topics.

To explain the normalization of  $E_A$ , assume that  $\mathcal{X}(\alpha)$  is the union of  $N(\alpha)$  disjoint balls of radius  $\alpha$ , where the volume of  $\mathcal{X}(\alpha)$  is proportional to  $N(\alpha)\alpha^{n-1}$ . Then  $E_A(\mathcal{X}, \alpha)$  is proportional to the number of combinations of four such balls such that the sum of the centers of the first two balls is approximately equal to the sum of the centers of the other two.

Motivated by (1.5), we assume that  $N(\alpha) \sim \alpha^{-\delta}$ . Then

$$\alpha^{-2\delta} \lesssim E_A(\mathcal{X}, \alpha) \lesssim \alpha^{-3\delta}. \quad (1.18)$$

Indeed, the upper bound follows from the fact that the first three balls determine the fourth one uniquely, and the lower bound follows from considering combinations of the form  $(\eta_1, \eta_1, \eta_3, \eta_3)$ .

We will use the additive energy of the images of the limit set  $\Lambda_\Gamma$  by the map

$$\mathcal{G}(y, y') = \frac{y' - (y \cdot y')y}{1 - y \cdot y'} \in \mathbb{R}^n, \quad y, y' \in \mathbb{S}^{n-1} \subset \mathbb{R}^n, \quad y \neq y', \quad (1.19)$$

which is half the stereographic projection of  $y'$  with the base point  $y$  – see Figure 3. We have  $\mathcal{G}(y, y') \perp y$ , therefore we may think of it as a vector in  $T_y \mathbb{S}^{n-1}$ , or (pairing with the round metric on the sphere) as a vector in  $T_y^* \mathbb{S}^{n-1}$ . Note that  $\mathcal{G}$  is related to the leaves of (1.9) since

$$\partial_y \log(|y - y'|^2) = -\mathcal{G}(y, y'). \quad (1.20)$$

**Definition 1.3.** We say that  $\Lambda_\Gamma$  satisfies the **additive energy bound** with exponent  $\beta_E > 0$ , if for each  $C_1 > 0$  there exists  $C > 0$  such that for all  $\alpha \in (0, 1)$ ,

$$\sup_{y_0 \in \Lambda_\Gamma} E_A(\mathcal{G}(y_0, \Lambda_\Gamma) \cap B(0, C_1), \alpha) \leq C\alpha^{-3\delta + \beta_E}. \quad (1.21)$$

One can also interpret the sets  $\mathcal{G}(y_0, \Lambda_\Gamma)$  in terms of the dynamics of the geodesic flow on  $M$  using horocyclic flows – see (7.4) and (7.5).

Given an additive energy bound, we obtain a fractal uncertainty principle and thus (by Theorem 3) an essential spectral gap:

**Theorem 4.** *Assume that  $\Lambda_\Gamma$  satisfies the additive energy bound with exponent  $\beta_E > 0$ . Then  $\Lambda_\Gamma$  satisfies the fractal uncertainty principle with exponent*

$$\beta = \frac{3}{8} \left( \frac{n-1}{2} - \delta \right) + \frac{\beta_E}{16}. \quad (1.22)$$

**Remark.** Note that by (1.18), the maximal  $\beta_E$  for which Definition 1.3 may hold is  $\beta_E = \delta$ . Plugged into (1.22), this gives an essential spectral gap of size  $\frac{3(n-1)-5\delta}{16}$ , which improves over (1.2) only when  $\delta \in (\frac{5}{11}(n-1), \frac{3}{5}(n-1))$ .

Theorem 4 is proved using an  $L^4$  estimate on the Fourier transforms of  $\mathbb{1}_{\mathcal{G}(y_0, \Lambda_\Gamma(h^{\rho/2}))}$  for  $y_0 \in \Lambda_\Gamma$  obtained from the additive energy bound. Here we have to replace the original  $h^\rho$  neighborhood of  $\Lambda_\Gamma$  by a bigger  $h^{\rho/2}$  neighborhood to approximate correlations between different leaves of (1.9) restricted to  $\Lambda_\Gamma(h^{\rho/2})$  using the Fourier transform. Roughly speaking, the leaves which are farther than  $h^{1/2}$  apart have an  $\mathcal{O}(h^\infty)$  correlation and for the leaves which are closer than  $h^{1/2}$  to each other, the difference of the phase functions in the resulting integral can be well approximated by its linear part – see the paragraph following (5.16). The enlargement of the neighborhood to  $\Lambda_\Gamma(h^{\rho/2})$  causes the loss of a factor of  $\frac{1}{2}$  in the size of the gap; together with a factor of  $\frac{3}{4}$  coming from the use of the  $L^4$  bound (rather than  $L^\infty$ ) this explains the factor of  $\frac{3}{8}$  in (1.22).

**1.3. Additive energy via Ahlfors-David regularity.** We now restrict to dimension  $n = 2$  and show that the limit set  $\Lambda_\Gamma \subset \mathbb{S}^1$  of a convex co-compact Fuchsian group  $\Gamma$  with  $\delta \in (0, 1)$  satisfies the additive energy bound with some positive exponent. For that we use the following regularity property:

**Definition 1.4.** *Let  $(\mathcal{M}, d)$  be a complete metric space with more than one element. We say a closed set  $\mathcal{X} \subset \mathcal{M}$  is  $\delta$ -regular with constant  $C_\mathcal{X}$  if for all  $x \in \mathcal{X}$  we have*

$$C_\mathcal{X}^{-1} r^\delta \leq \mu_\delta(\mathcal{X} \cap B(x, r)) \leq C_\mathcal{X} r^\delta, \quad 0 < r < \text{diam}(\mathcal{M}) \quad (1.23)$$

where  $B(x, r)$  is the metric ball of radius  $r$  centered at  $x$  and  $\mu_\delta$  is the  $\delta$ -dimensional Hausdorff measure.

Sets with this property are also known as *Ahlfors-David regular*. See [DaSe] for an introduction to  $\delta$ -regular sets. While Definition 1.4 is phrased using  $\delta$ -dimensional Hausdorff measure, any other Borel outer measure could be used instead (in particular, for limit sets of convex co-compact Fuchsian groups the Patterson–Sullivan measure could be used). This is discussed further in Lemma 7.5 below.

The limit set  $\Lambda_\Gamma \subset \mathbb{S}^{n-1}$  of a convex co-compact Fuchsian group  $\Gamma$  is  $\delta$ -regular with  $\delta$  defined in (5.2) – see [Su, Theorem 7] and [Bo07, Lemma 14.13 and Theorem 14.14].

We denote the associated regularity constant by

$$\mathbf{C} := C_{\Lambda_\Gamma}. \quad (1.24)$$

Using  $\delta$ -regularity of  $\Lambda_\Gamma$ , we obtain the following additive energy bound. Combined with Theorems 3 and 4, it implies Theorem 2 and thus Theorem 1.

**Theorem 5.** *Let  $M = \Gamma \backslash \mathbb{H}^2$  be a convex co-compact hyperbolic surface with limit set  $\Lambda_\Gamma \subset \mathbb{S}^1$  of dimension  $\delta \in (0, 1)$ . Then  $\Lambda_\Gamma$  satisfies the additive energy bound in the sense of Definition 1.3 with exponent*

$$\beta_E := \delta \exp \left[ -\mathbf{K}(1 - \delta)^{-28}(1 + \log^{14} \mathbf{C}) \right], \quad (1.25)$$

where  $\mathbf{C}$  is defined in (1.24) and  $\mathbf{K}$  is a global constant.

**Remarks.** (i) The specifics of the bound (1.25) are not particularly important. The key point is that the exponent  $\beta_E$  in (1.21) is independent of  $\alpha$ , and it can be computed explicitly. We did not compute the value of  $\mathbf{K}$ , but in principle it can be done without much difficulty.

(ii) In dimensions  $n > 2$ , Theorem 5 no longer holds in general as shown by the example of the hyperbolic cylinder in three dimensions (see for instance [DaDy, Appendix A]). In this example, the limit set  $\Lambda_\Gamma$  is a great circle on  $\mathbb{S}^2$ , and the stereographic projections  $\mathcal{G}(y_0, \Lambda_\Gamma)$  are straight lines, which saturate the upper bound in (1.18). See §6.8.2 for possible generalizations to higher dimensions.

Theorem 5 follows from a general result bounding additive energy of Ahlfors-David regular sets, stated as Theorem 6 in §6; the proof of Theorem 6 can schematically be explained as follows (see §6.1 for more details):

- (1) Ahlfors-David regular sets cannot contain large subsets of arithmetic progressions. This follows by a direct argument using (1.23) and the fact that  $\delta < 1$ .
- (2) A variant of Freĭman's theorem from additive combinatorics asserts that any set with large additive energy must contain large subsets of generalized arithmetic progressions. Together with (1) this implies that Ahlfors-David regular sets cannot have extremely large (i.e. near maximal) additive energy.
- (3) Ahlfors-David regular sets also have a certain type of coarse self-similarity. This allows us to analyze them at many scales and at many different locations. Since Ahlfors-David regular sets cannot have extremely large additive energy at any scale or at any location, we can perform a multi-scale analysis to conclude that such sets must actually have small additive energy.

#### 1.4. Structure of the paper.

- In §2, we review certain notions in semiclassical analysis, in particular pseudo-differential and Fourier integral operators.

- In §3, we study an anisotropic pseudodifferential calculus associated to a Lagrangian foliation.
- In §4, we study geometric and dynamical properties of hyperbolic manifolds and, using the calculus of §3, prove Theorem 3.
- In §5, we discuss the fractal uncertainty principle and prove Theorem 4.
- In §6, we prove that Ahlfors-David regular sets have small additive energy.
- In §7, we establish Ahlfors-David regularity of the stereographic projections of the limit set and prove Theorem 5. We also obtain locally uniform bounds on the regularity constant for 3-funneled surfaces.
- In Appendix A, we prove several technical lemmas used in §4 and §7.

## 2. SEMICLASSICAL PRELIMINARIES

In this section, we give a brief review of semiclassical analysis. For a comprehensive introduction to the subject, the reader is referred to [Zw]. We partially follow the presentation of [DyZw, Appendix E] and [DyGu14a, DaDy, Dy].

**2.1. Pseudodifferential operators.** Let  $M$  be a manifold. For  $k \in \mathbb{R}$ , we say that  $a(x, \xi) \in C^\infty(T^*M)$  lies in the symbol class  $S_{1,0}^k(T^*M)$  if it satisfies the derivative bounds

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \leq C_{\alpha\beta K} \langle \xi \rangle^{k-|\beta|}, \quad x \in K, \quad (2.1)$$

for each compact set  $K \subset M$ . We restrict ourselves to the subset of *polyhomogeneous, or classical, symbols*  $S^k(T^*M) \subset S_{1,0}^k(T^*M)$  which have asymptotic expansions  $a(x, \xi) \sim \sum_{j=0}^{\infty} a_j(x, \xi)$  as  $|\xi| \rightarrow \infty$  where each  $a_j$  is positively homogeneous in  $\xi$  of degree  $k - j$ .

A family of symbols  $b(x, \xi; h) \in S_{1,0}^k(T^*M)$  depending on a small parameter  $h > 0$  is said to lie in the class  $S_h^k(T^*M)$  if it has the following expansion as  $h \rightarrow 0$ :

$$b(x, \xi; h) \sim \sum_{\ell=0}^{\infty} h^\ell a_\ell(x, \xi), \quad a_\ell \in S^{k-\ell}(T^*M). \quad (2.2)$$

See for instance [DyZw, §E.1.2] and [Va13b, §2] for details.

If  $a \in S_{1,0}^k(T^*\mathbb{R}^n)$  satisfies (2.1) uniformly in  $x \in \mathbb{R}^n$ , then we can quantize it by the following formula (see [Zw, §4.1.1] and [DyZw, §E.1.4])

$$\text{Op}_h(a)f(x) = (2\pi h)^{-n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}(x-y)\cdot\xi} a(x, \xi) f(y) dy d\xi, \quad (2.3)$$

which gives an operator  $\text{Op}_h(a)$  acting on the space  $\mathcal{S}(\mathbb{R}^n)$  of Schwartz functions, as well as on the dual space  $\mathcal{S}'(\mathbb{R}^n)$  of tempered distributions.

Following [DyZw, §E.1.5] and [Zw, §14.2.2], for a general manifold  $M$  we consider the class  $\Psi_h^k(M)$  of semiclassical pseudodifferential operators with symbols in  $S_h^k(T^*M)$ .

We denote by

$$\sigma_h : \Psi_h^k(M) \rightarrow S^k(T^*M)$$

the principal symbol map. Operators in  $\Psi_h^k$  act on semiclassical Sobolev spaces  $H_{h,\text{comp}}^s \rightarrow H_{h,\text{loc}}^{s-k}$ , see [DyZw, §E.1.6] and [Zw, §14.2.4]. We will often use the class  $\Psi_h^{\text{comp}}(M)$  of operators whose full symbols are essentially compactly supported in  $T^*M$  and whose Schwartz kernels are compactly supported in  $M \times M$ .

For  $A \in \Psi_h^k(M)$ , denote by  $\text{WF}_h(A)$  its semiclassical wavefront set, which is the essential support of its full symbol – see for instance [DyZw, §E.2.1] and [DyZw16, Appendix C.1]. Then  $\text{WF}_h(A)$  is a closed subset of the fiber-radially compactified cotangent bundle  $\overline{T^*}M \supset T^*M$ , see for instance [DyZw, §E.1.2] or [Va13b, §2]. For  $A, B \in \Psi_h^k(M)$  and an open set  $U \subset \overline{T^*}M$ , we say that

$$A = B + \mathcal{O}(h^\infty) \quad \text{microlocally in } U,$$

if  $\text{WF}_h(A - B) \cap U = \emptyset$ . We also use the notion of wavefront sets of  $h$ -tempered distributions and operators, see for instance [DyZw, §E.2.3] or [DyZw16, §2.3].

Let  $B = B(h) : \mathcal{D}'(M) \rightarrow C_0^\infty(M)$  be an  $h$ -tempered family of smoothing operators and assume that the wavefront set  $\text{WF}'_h(B) \subset \overline{T^*}(M \times M)$  is a compact subset of  $T^*(M \times M)$ . We say that  $B$  is *pseudolocal* if  $\text{WF}'_h(B)$  is contained in the diagonal  $\Delta(T^*M) \subset T^*(M \times M)$ . For a pseudolocal operator  $B$ , we consider the set  $\text{WF}_h(B) \subset T^*M$  defined by

$$\text{WF}'_h(B) = \{(x, \xi, x, \xi) \mid (x, \xi) \in \text{WF}_h(B)\}. \quad (2.4)$$

Note that operators in  $\Psi_h^{\text{comp}}(M)$  are pseudolocal and their definition of wavefront set given in [DyZw, §E.2.1] agrees with the one given by (2.4).

**2.2. Fourier integral operators.** We next introduce semiclassical Fourier integral operators. Let  $\varkappa : U_2 \rightarrow U_1$  be a canonical transformation (that is, a symplectomorphism), where  $U_j \subset T^*M_j$  are open sets and  $M_j$  are manifolds of the same dimension. Define the graph of  $\varkappa$  by

$$\text{Gr}(\varkappa) := \{(x, \xi, y, \eta) \mid (y, \eta) \in U_2, (x, \xi) = \varkappa(y, \eta)\} \subset T^*(M_1 \times M_2). \quad (2.5)$$

Let  $\xi dx$  and  $\eta dy$  be the canonical 1-forms on  $T^*U_1$  and  $T^*U_2$  respectively. Since  $\varkappa$  is a canonical transformation, the restriction  $(\xi dx - \eta dy)|_{\text{Gr}(\varkappa)}$  is a closed 1-form. We require that  $\varkappa$  is *exact* in the sense that this restriction is an exact form, and fix an antiderivative

$$F \in C^\infty(\text{Gr}(\varkappa)), \quad (\xi dx - \eta dy)|_{\text{Gr}(\varkappa)} = dF. \quad (2.6)$$

For a canonical transformation  $\varkappa$  with a fixed antiderivative  $F$ , we consider the class  $I_h^{\text{comp}}(\varkappa)$  of compactly supported and microlocalized Fourier integral operators associated to  $\varkappa$  – see for instance [GuSt77, Chapter 5], [GuSt13, Chapter 8], [DyGu14a,

§3.2], [DaDy, §3.2],<sup>†</sup> [Dy, §3.2], and the references there. We adopt a convention that operators in  $I_h^{\text{comp}}(\varkappa)$  act  $\mathcal{D}'(M_2) \rightarrow C_0^\infty(M_1)$ .

We list some basic properties of the class  $I_h^{\text{comp}}(\varkappa)$ :

- each  $B \in I_h^{\text{comp}}(\varkappa)$  is bounded uniformly in  $h$  on the spaces  $H_{h,\text{loc}}^s(M_2) \rightarrow H_{h,\text{comp}}^{s'}(M_1)$  for all  $s, s' \in \mathbb{R}$ , and  $\text{WF}'_h(B) \subset \text{Gr}(\varkappa)$ ;
- if  $\varkappa : U_2 \rightarrow U_1$ ,  $U' \subset U_2$ , and  $\varkappa' := \varkappa|_{U'}$ , then  $B \in I_h^{\text{comp}}(\varkappa')$  if and only if  $B \in I_h^{\text{comp}}(\varkappa)$  and  $\text{WF}'_h(B) \subset \text{Gr}(\varkappa')$ ;
- if  $\varkappa : T^*M \rightarrow T^*M$  is the identity map with the zero antiderivative, then  $B \in I_h^{\text{comp}}(\varkappa)$  if and only if  $B \in \Psi_h^{\text{comp}}(M)$ ;
- if  $B \in I_h^{\text{comp}}(\varkappa)$ , then  $B^* \in I_h^{\text{comp}}(\varkappa^{-1})$ , with the antiderivatives on  $\text{Gr}(\varkappa)$  and  $\text{Gr}(\varkappa^{-1})$  summing up to zero;
- if  $\varkappa : U_2 \rightarrow U_1$ ,  $\varkappa' : U_3 \rightarrow U_2$ , and  $B \in I_h^{\text{comp}}(\varkappa)$ ,  $B' \in I_h^{\text{comp}}(\varkappa')$ , then  $BB' \in I_h^{\text{comp}}(\varkappa \circ \varkappa')$ , with the antiderivative on  $\text{Gr}(\varkappa \circ \varkappa')$  chosen as the sum of the antiderivatives on  $\text{Gr}(\varkappa)$  and  $\text{Gr}(\varkappa')$ .

To give a concrete expression for elements of  $I_h^{\text{comp}}(\varkappa)$ , assume that  $\varkappa$  is parametrized by a nondegenerate phase function  $\Phi(x, y, \zeta) \in C^\infty(U_\Phi; \mathbb{R})$ ,  $U_\Phi \subset M_1 \times M_2 \times \mathbb{R}^m$ , in the sense that the differentials  $d(\partial_{\zeta_1} \Phi), \dots, d(\partial_{\zeta_m} \Phi)$  are independent on the critical set

$$\mathcal{C}_\Phi = \{(x, y, \zeta) \in U_\Phi \mid \partial_\zeta \Phi(x, y, \zeta) = 0\}$$

and the graph  $\text{Gr}(\varkappa)$  is given by

$$\text{Gr}(\varkappa) = j_\Phi(\mathcal{C}_\Phi), \quad j_\Phi : (x, y, \zeta) \mapsto (x, \partial_x \Phi(x, y, \zeta), y, -\partial_y \Phi(x, y, \zeta)). \quad (2.7)$$

The corresponding antiderivative is just the pullback of  $\Phi$  from  $\mathcal{C}_\Phi$  to  $\text{Gr}(\varkappa)$  by the map  $j_\Phi$ . Then any operator  $B \in I_h^{\text{comp}}(\varkappa)$  has the following form modulo  $\mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ :

$$Bf(x) = (2\pi h)^{-\frac{m+n}{2}} \int_{M_1 \times \mathbb{R}^m} e^{\frac{i}{h} \Phi(x, y, \zeta)} b(x, y, \zeta; h) dy d\zeta \quad (2.8)$$

where  $n = \dim M_1 = \dim M_2$  and  $b$  is a compactly supported symbol on  $U_\Phi$ , that is an  $h$ -dependent family of smooth functions with support contained in some  $h$ -independent compact set which has an asymptotic expansion in nonnegative integer powers of  $h$ . Moreover, local principal symbol calculus shows that

$$b(x, y, \zeta; 0) = 0 \quad \text{for all } (x, y, \zeta) \in \mathcal{C}_\Phi \implies B \in hI_h^{\text{comp}}(\varkappa). \quad (2.9)$$

See for example [DyGu14a, §3.2] for details.

A special case is when  $M_2$  is an open subset of  $\mathbb{R}^n$  and  $\text{Gr}(\varkappa)$  projects diffeomorphically onto the  $(x, \eta)$  variables. Let  $F \in C^\infty(\text{Gr}(\varkappa))$  be the fixed antiderivative, and

<sup>†</sup>The presentation in [DaDy, §3.2] contained an error because the Fourier integral operators associated to the identity map were not necessarily pseudodifferential operators with classical symbols due to a possible constant phase factor  $e^{ic/h}$ . We correct it here by fixing the antiderivative, which is always possible locally.

define the *generating function*  $S(x, \eta) \in C^\infty(U_S; \mathbb{R})$  by the formula  $S(x, \eta) = F + y \cdot \eta$ , where  $\text{Gr}(\varkappa)$  is parametrized by  $(x, \eta) \in U_S \subset M_1 \times \mathbb{R}^n$ . Then

$$\text{Gr}(\varkappa) = \{\xi = \partial_x S(x, \eta), y = \partial_\eta S(x, \eta), (x, \eta) \in U_S\} \quad (2.10)$$

implying that  $\varkappa$  is parametrized in the sense of (2.7) by the function  $(x, y, \zeta) \mapsto S(x, \zeta) - y \cdot \zeta$ . Each  $B \in I_h^{\text{comp}}(\varkappa)$  has the following form modulo  $\mathcal{O}(h^\infty)_{\mathcal{D}'(M_2) \rightarrow C_0^\infty(M_1)}$ :

$$Bf(x) = (2\pi h)^{-n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}(S(x, \eta) - y \cdot \eta)} b(x, \eta; h) \chi(y) f(y) dy d\eta, \quad f \in \mathcal{D}'(M_2), \quad (2.11)$$

where  $n = \dim M_j$ ,  $b(x, \eta; h)$  is a compactly supported symbol on  $U_S$ , and  $\chi \in C_0^\infty(M_2)$  is any function such that  $\chi = 1$  near  $\partial_\eta S(\text{supp } b)$ . (The resulting operator is independent of the choice of  $\chi$  modulo  $\mathcal{O}(h^\infty)_{\mathcal{D}'(M_2) \rightarrow C_0^\infty(M_1)}$ .)

As remarked in [DaDy, §3.2],  $\varkappa$  can locally be written in the form (2.10) for some choice of local coordinates on  $M_2$  as long as its domain does not intersect the zero section of  $T^*M_2$ . The latter condition can be arranged locally by composing  $\varkappa$  with a transformation of the form  $(y, \eta) \mapsto (y, \eta - d\psi(y))$  for some  $\psi \in C^\infty(M_2)$ , which amounts to multiplying the resulting operators by  $e^{i\psi(y)/h}$  – see Lemma 2.1 below.

We next discuss microlocal inverses of Fourier integral operators. Assume that  $B \in I_h^{\text{comp}}(\varkappa)$ ,  $B' \in I_h^{\text{comp}}(\varkappa^{-1})$ . Then  $BB' \in \Psi_h^{\text{comp}}(M_1)$ ,  $B'B \in \Psi_h^{\text{comp}}(M_2)$ ,  $\text{WF}_h(BB') \subset U_1$ ,  $\text{WF}_h(B'B) \subset U_2$ , and (as shown in the case of (2.11) by an explicit application of the method of stationary phase and in general is a form of Egorov's Theorem)

$$\sigma_h(B'B) = \sigma_h(BB') \circ \varkappa. \quad (2.12)$$

We call  $B \in I_h^{\text{comp}}(\varkappa)$  *elliptic* at a point  $(x, \xi, y, \eta) \in \text{Gr}(\varkappa)$ , if there exists  $B' \in I_h^{\text{comp}}(\varkappa^{-1})$  such that  $\sigma_h(BB')(x, \xi) \neq 0$  (in fact, this is equivalent to requiring that  $\sigma_h(BB^*)(x, \xi) \neq 0$ ). For  $B$  given by (2.8), this simply means that  $b(x, y, \zeta; 0) \neq 0$  where  $(x, y, \zeta) = j_\Phi^{-1}(x, \xi, y, \eta) \in \mathcal{C}_\Phi$ . For each point in  $\text{Gr}(\varkappa)$ , there exist operators in  $I_h^{\text{comp}}(\varkappa)$  elliptic at this point.

If  $V_j \subset U_j$  are compact subsets such that  $\varkappa(V_2) = V_1$ , then we say that  $B, B'$  *quantize*  $\varkappa$  near  $V_1 \times V_2$  if

$$\begin{aligned} BB' &= 1 + \mathcal{O}(h^\infty) \quad \text{microlocally near } V_1, \\ B'B &= 1 + \mathcal{O}(h^\infty) \quad \text{microlocally near } V_2. \end{aligned} \quad (2.13)$$

Such operators  $B, B'$  exist if  $V_2 = \{(y, \eta)\}$  for any given point  $(y, \eta) \in U_2$  (and thus if  $V_2$  is a sufficiently small neighborhood of  $(y, \eta)$ ). To show this, take  $B \in I_h^{\text{comp}}(\varkappa)$  elliptic at  $(\varkappa(y, \eta), y, \eta)$  and  $B'_0 \in I_h^{\text{comp}}(\varkappa^{-1})$  such that  $\sigma_h(BB'_0) \neq 0$  on  $V_1$ . Multiplying  $B'_0$  on the right by an elliptic parametrix of  $BB'_0$  (see for instance [DyZw, §E.2.2] and [DyZw16, Proposition 2.4]), we obtain  $B' \in I_h^{\text{comp}}(\varkappa)$  such that  $BB' = 1 + \mathcal{O}(h^\infty)$  microlocally near  $V_1$ . By (2.12), we have  $\sigma_h(B'_0 B) \neq 0$  on  $V_2$ , so we can construct

$B'' \in I_h^{\text{comp}}(\varkappa)$  such that  $B''B = 1 + \mathcal{O}(h^\infty)$  microlocally near  $V_2$ . Then

$$\text{WF}'_h(B' - B'') \cap (V_1 \times V_2) \subset \text{WF}'_h((B''B)B' - B''(BB')) = \emptyset,$$

therefore (2.13) holds. One could also define  $B, B'$  as solutions of an evolution equation, see [Zw, Theorem 11.5] and [DaDy, §3.2].

One useful family of Fourier integral operators is given by the following

**Lemma 2.1.** *Let  $\varphi : M_1 \rightarrow M_2$  be a diffeomorphism and  $\psi \in C^\infty(M_1)$ . Consider the operator*

$$B = B(h) : \mathcal{D}'(M_2) \rightarrow \mathcal{D}'(M_1), \quad Bf(x) = e^{i\psi(x)/h} f(\varphi(x)).$$

Then for each  $A_j \in \Psi_h^{\text{comp}}(M_j)$ , we have  $A_1B, BA_2 \in I_h^{\text{comp}}(\varkappa^{-1})$ , where

$$\varkappa : T^*M_1 \rightarrow T^*M_2, \quad \varkappa(x, \xi) = (\varphi(x), (d\varphi(x))^{-T} \cdot (\xi - d\psi(x))),$$

and the antiderivative is given by  $\psi(x)$ .

*Proof.* It suffices to consider the case when  $M_1, M_2$  are open subsets of  $\mathbb{R}^n$ . Let  $A_2 = \text{Op}_h(a)\chi$ , where  $a(y, \eta; h)$  is compactly supported in  $M_2 \times \mathbb{R}^n$  and  $\chi \in C_0^\infty(M_2)$  is equal to 1 near the projection of  $\text{supp } a$ . Then

$$BA_2f(x) = (2\pi h)^{-n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}((\varphi(x)-y)\cdot\eta+\psi(x))} a(\varphi(x), \eta; h)\chi(y)f(y) dyd\eta.$$

This has the form (2.11) with

$$S(x, \eta) = \varphi(x) \cdot \eta + \psi(x), \quad b(x, \eta; h) = a(\varphi(x), \eta; h),$$

and it is straightforward to see that  $\varkappa^{-1}$  is given by (2.10). The case of  $A_1B$  is reduced to the case of  $BA_2$  by considering adjoint operators.  $\square$

### 3. CALCULUS ASSOCIATED TO A LAGRANGIAN FOLIATION

In this section, we define a class of exotic pseudodifferential operators associated to a Lagrangian foliation. The symbols of these operators are allowed to vary on the constant scale along the foliation and on the scale  $h^\rho$ ,  $0 \leq \rho < 1$ , in the directions transversal to the foliation. For  $\rho > \frac{1}{2}$ , the resulting operators will not generally lie in the exotic calculus  $\Psi_{1/2}$  (see for instance [DaDy, §5.1]), yet they form an algebra with properties similar to those of standard pseudodifferential operators.

A similar (in fact, sharper in certain ways as it allowed for  $\rho = 1$  and  $\Psi_{1/2}$  behavior in some directions) second microlocal calculus associated to a hypersurface has previously been developed by Sjöstrand–Zworski [SjZw, §5]; for a calculus associated to a Lagrangian submanifold in the analytic category, see [De, Chapter 2] and the references given there.

**3.1. Foliations and symbols.** We start with the definition of a Lagrangian foliation:

**Definition 3.1.** *Let  $M$  be a manifold,  $U \subset T^*M$  be an open set, and*

$$L_{(x,\xi)} \subset T_{(x,\xi)}(T^*M), \quad (x, \xi) \in U$$

*a family of subspaces depending smoothly on  $(x, \xi)$ . We say that  $L$  is a **Lagrangian foliation** on  $U$  if*

- $L_{(x,\xi)}$  is integrable in the sense that if  $X, Y$  are two vector fields on  $U$  lying in  $L$  at each point (we denote this by  $X, Y \in C^\infty(U; L)$ ), then the Lie bracket  $[X, Y]$  lies in  $C^\infty(U; L)$  as well;
- $L_{(x,\xi)}$  is a Lagrangian subspace of  $T_{(x,\xi)}(T^*M)$  for each  $(x, \xi) \in U$ .

Another way to think about a Lagrangian foliation is in terms of its leaves, which are Lagrangian submanifolds whose tangent spaces are given by  $L$ . The existence of these leaves follows from Frobenius's Theorem, see Lemma 3.6 below.

We consider the following class of symbols:

**Definition 3.2.** *Let  $L$  be a Lagrangian foliation on  $U \subset T^*M$ , and fix  $\rho \in [0, 1)$ . We say that a function  $a(x, \xi; h)$  is a (compactly supported) symbol of class  $S_\rho$  with respect to  $L$ , and write*

$$a \in S_{L,\rho}^{\text{comp}}(U),$$

*if for each  $h \in (0, h_0)$ ,  $(x, \xi) \mapsto a(x, \xi; h)$  is a smooth function on  $U$  supported inside some  $h$ -independent compact set and it satisfies the derivative bounds (with the constant  $C$  depending on  $Y_j, Z_j$ , but not on  $h$ )*

$$\sup_{x,\xi} |Y_1 \dots Y_m Z_1 \dots Z_k a(x, \xi; h)| \leq Ch^{-\rho k}, \quad (3.1)$$

*for each vector fields  $Y_1, \dots, Y_m, Z_1, \dots, Z_k$  on  $U$  such that  $Y_1, \dots, Y_m \in C^\infty(U; L)$ .*

The following statement is useful for constructing symbols in the class  $S_{L,\rho}^{\text{comp}}$ :

**Lemma 3.3.** *Let  $M_1$  be a compact manifold and  $V_0(h) \subset V_1(h) \subset M_1$  be  $h$ -dependent sets satisfying*

$$d(V_0(h), M_1 \setminus V_1(h)) > \varepsilon h^\rho$$

*for some fixed  $\varepsilon > 0, \rho \in [0, 1)$  and all  $h \in (0, 1)$ . Then there exists  $\chi(h) \in C_0^\infty(M_1; [0, 1])$  such that for all  $h \in (0, 1)$ ,*

$$\text{supp}(1 - \chi(h)) \cap V_0(h) = \emptyset, \quad \text{supp} \chi(h) \subset V_1(h); \quad (3.2)$$

$$\sup_{M_1} |\partial^\alpha \chi| \leq C_\alpha h^{-\rho|\alpha|}. \quad (3.3)$$

*Proof.* By a partition of unity we reduce to the case when  $V_1(h)$  is contained in a small coordinate neighborhood on  $M_1$ ; therefore, it suffices to consider the case  $M_1 = \mathbb{R}^n$ . Let  $d(\cdot, \cdot)$  be the Euclidean distance function. Put

$$V_2(h) := \{x \in \mathbb{R}^n \mid d(x, V_0(h)) \leq \varepsilon h^\rho/2\},$$

then (here  $B(x, r)$  denotes the ball of radius  $r$  centered at  $x$ )

$$\begin{aligned} x \in V_0(h) &\implies B(x, \varepsilon h^\rho/2) \subset V_2(h), \\ x \in V_2(h) &\implies B(x, \varepsilon h^\rho/2) \subset V_1(h). \end{aligned} \tag{3.4}$$

Take nonnegative  $\psi \in C_0^\infty(B(0, \varepsilon/2))$  such that  $\int \psi = 1$ , and put (here  $m = \dim M_1$ )

$$\chi(x; h) := h^{-m\rho} \int_{V_2(h)} \psi\left(\frac{x-y}{h^\rho}\right) dy.$$

It follows immediately from (3.4) that  $\chi$  satisfies (3.2). Moreover, by putting derivatives on  $\psi$  we obtain the derivative bounds (3.3), finishing the proof.  $\square$

To keep track of the essential supports of symbols in  $S_{L,\rho}^{\text{comp}}(U)$  in an  $h$ -dependent way, we use the following

**Definition 3.4.** *Assume that  $a(x, \xi; h)$  is an  $h$ -dependent family of smooth functions in  $(x, \xi) \in U$ , and  $h_j \rightarrow 0$ ,  $(x_j, \xi_j) \in U$  are some sequences. We say that  $a$  is  $\mathcal{O}(h^\infty)$  along the sequence  $(x_j, \xi_j, h_j)$ , if for each  $N$  and each vector fields  $Z_1, \dots, Z_L$  on  $U$ , there exists a constant  $C$  such that*

$$|Z_1 \dots Z_L a(x_j, \xi_j; h_j)| \leq C h_j^N.$$

We next introduce local canonical coordinates bringing an arbitrary Lagrangian foliation to a normal form. Let  $L_0$  be the Lagrangian foliation on  $T^*\mathbb{R}^n$  given by the fibers of the cotangent bundle; that is, in the standard coordinates  $(y, \eta)$  on  $T^*\mathbb{R}^n$ ,

$$L_0 = \text{span}(\partial_{\eta_1}, \dots, \partial_{\eta_n})$$

is the annihilator of  $dy$ .

**Definition 3.5.** *Let  $L$  be a Lagrangian foliation on  $U \subset T^*M$ . A **Lagrangian chart** is a symplectomorphism*

$$\varkappa : U_0 \rightarrow V, \quad U_0 \subset U, \quad V \subset T^*\mathbb{R}^n,$$

such that  $d\varkappa(x, \xi) \cdot L_{(x,\xi)} = (L_0)_{\varkappa(x,\xi)}$  for each  $(x, \xi) \in U_0$ .

The basic properties of Lagrangian charts are given by

**Lemma 3.6.** *1. Let  $L$  be a Lagrangian foliation on  $U \subset T^*M$  and  $(x_0, \xi_0) \in U$ . Then there exists a Lagrangian chart  $\varkappa : U_0 \rightarrow T^*\mathbb{R}^n$  on some neighborhood  $U_0 \subset U$  of  $(x_0, \xi_0)$ .*

2. Assume that  $\varkappa : V \rightarrow V'$ , where  $V, V' \subset T^*\mathbb{R}^n$  are open, is a symplectomorphism which preserves the foliation  $L_0$ , and  $(y_0, \eta_0) \in V$ . Then there exists  $\varepsilon > 0$  such that

$$\varkappa(y, \eta) = (\varphi(y), (d\varphi(y))^{-T} \cdot (\eta - d\psi(y))), \quad (y, \eta) \in B(y_0, \varepsilon) \times B(\eta_0, \varepsilon), \quad (3.5)$$

for some diffeomorphism  $\varphi : B(y_0, \varepsilon) \rightarrow \mathbb{R}^n$  onto its image and some function  $\psi \in C^\infty(B(y_0, \varepsilon); \mathbb{R})$ .

*Proof.* 1. Since  $L$  is integrable, by Frobenius's Theorem [HöIII, Theorem C.1.1] there exist local coordinates  $(y, \tilde{\eta})$  in a neighborhood of  $(x_0, \xi_0)$  such that  $L$  is the annihilator of  $dy$ . Moreover, since  $L$  is Lagrangian, we have  $\{y_j, y_k\} = 0$ . Now, by Darboux Theorem [HöIII, Theorem 21.1.6] there exists a set of functions  $\eta_1, \dots, \eta_n$  defined near  $(x_0, \xi_0)$  such that

$$\{y_j, y_k\} = \{\eta_j, \eta_k\} = 0, \quad \{\eta_j, y_k\} = \delta_{jk}.$$

The map  $(x, \xi) \mapsto (y, \eta)$  is a Lagrangian chart in a neighborhood of  $(x_0, \xi_0)$ .

2. Define the functions  $y', \eta'$  on  $V$  by setting  $\varkappa : (y, \eta) \mapsto (y', \eta')$ . Since the annihilators of  $dy'$  and  $dy$  are the same (and both equal to  $L_0$ ), we have  $y' = \varphi(y)$  for  $(y, \eta) \in B(y_0, \varepsilon) \times B(\eta_0, \varepsilon)$ , some  $\varepsilon > 0$ , and some diffeomorphism onto its image  $\varphi : B(y_0, \varepsilon) \rightarrow \mathbb{R}^n$ .

Since  $\{y'_j, \eta'_k\} = \delta_{jk}$ , we have

$$\eta' = (d\varphi(y))^{-T} \cdot (\eta - F(y)), \quad (y, \eta) \in B(y_0, \varepsilon) \times B(\eta_0, \varepsilon),$$

for some smooth map  $F : B(y_0, \varepsilon) \rightarrow \mathbb{R}^n$ . Since  $\{\eta'_j, \eta'_k\} = 0$ , we have  $F(y) = d\psi(y)$  for some  $\psi : B(y_0, \varepsilon) \rightarrow \mathbb{R}$ .  $\square$

**3.2. Calculus on  $\mathbb{R}^n$ .** We next develop the calculus for the case  $L = L_0$ . We have  $a(y, \eta; h) \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  if and only if  $a$  is supported inside some  $h$ -independent compact set and satisfies the derivative bounds

$$\sup_{y, \eta} |\partial_y^\alpha \partial_\eta^\beta a(y, \eta; h)| \leq C_{\alpha\beta} h^{-\rho|\alpha|}. \quad (3.6)$$

We derive several basic properties of quantizations of symbols in  $S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  by the map  $\text{Op}_h$  defined in (2.3):

**Lemma 3.7.** *For  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ , the operator  $\text{Op}_h(a)$  is bounded on  $L^2(\mathbb{R}^n)$  uniformly in  $h$ .*

*Proof.* We introduce the unitary rescaling operator

$$T_\rho : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n), \quad T_\rho u(y) = h^{\rho/4} u(h^{\rho/2} y).$$

It suffices to estimate the  $L^2 \rightarrow L^2$  norm of

$$T_\rho \text{Op}_h(a) T_\rho^{-1} = \text{Op}_h(a_\rho), \quad a_\rho(\tilde{y}, \tilde{\eta}; h) := a(h^{\rho/2} \tilde{y}, h^{-\rho/2} \tilde{\eta}; h).$$

It follows from (3.6) that  $a_\rho \in S_{\rho/2}$ , where the classes  $S_\delta$ ,  $0 \leq \delta \leq 1/2$ , are defined in [Zw, (4.4.5)]. It remains to apply [Zw, Theorem 4.23(ii)].  $\square$

**Lemma 3.8.** *Let  $a, b \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ . Then:*

1. *We have*

$$\text{Op}_h(a) \text{Op}_h(b) = \text{Op}_h(a\#b) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2},$$

where  $a\#b \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  and for each  $N$ ,

$$a\#b(y, \eta; h) = \sum_{j=0}^{N-1} \frac{(-ih)^j}{j!} (\partial_\eta \cdot \partial_{y'})^j (a(y, \eta; h) b(y', \eta'; h)) \Big|_{\substack{y'=y \\ \eta'=\eta}} + \mathcal{O}(h^{(1-\rho)N})_{S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)}.$$

2. *We have*

$$\text{Op}_h(a)^* = \text{Op}_h(a^*) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2},$$

where  $a^* \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  and for each  $N$ ,

$$a^*(y, \eta; h) = \sum_{j=0}^{N-1} \frac{(-ih)^j}{j!} (\partial_\eta \cdot \partial_y)^j \overline{a(y, \eta; h)} + \mathcal{O}(h^{(1-\rho)N})_{S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)}.$$

*Proof.* It suffices to apply [Zw, Theorems 4.14 and 4.17] to the rescaled symbols  $a_\rho, b_\rho \in S_{\rho/2}$  introduced in the proof of Lemma 3.7. The resulting symbols are  $\mathcal{O}(h^\infty)$  outside of a compact set and thus can be cut off to compactly supported symbols.  $\square$

Lemma 3.8 (or rather its trivial extension to symbols which are not compactly supported) implies that

$$\text{Op}_h(b_1) \text{Op}_h(a) \text{Op}_h(b_2) = \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}, \quad \text{supp } b_1 \cap \text{supp } b_2 = \emptyset, \quad (3.7)$$

for each  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  and  $h$ -independent  $b_1, b_2 \in C^\infty(\mathbb{R}^{2n})$  with all derivatives uniformly bounded. This in turn implies that the operator  $\text{Op}_h(a)$  is pseudolocal and  $\text{WF}'_h(\text{Op}_h(a))$  is compactly contained in  $T^*(\mathbb{R}^n \times \mathbb{R}^n)$ .

The next two lemmas establish invariance of the class of operators of the form  $\text{Op}_h(a)$ ,  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ , under conjugation by Fourier integral operators whose canonical transformations preserve the foliation  $L_0$ :

**Lemma 3.9.** *Assume that  $\varphi : V_1 \rightarrow V_2$  is a diffeomorphism, where  $V_j \subset \mathbb{R}^n$  are open sets,  $\psi \in C^\infty(V_1)$ , and  $\chi \in C_0^\infty(V_1)$ . Define the operators  $B : C^\infty(V_1) \rightarrow C_0^\infty(V_2)$ ,  $B' : C^\infty(V_2) \rightarrow C_0^\infty(V_1)$  by*

$$Bg(y') = e^{-i\psi(\varphi^{-1}(y'))/h} \chi(\varphi^{-1}(y')) f(\varphi^{-1}(y')), \quad B'f(y) = e^{i\psi(y)/h} \chi(y) f(\varphi(y)).$$

Then for each  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ ,

$$B' \text{Op}_h(a) B = \text{Op}_h(\tilde{a}) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2},$$

for some  $\tilde{a} \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  such that for each  $N$ ,

$$\tilde{a}(y, \eta; h) = \sum_{j=0}^{N-1} h^j L_j(\chi(y)\chi(y')a(\varphi(y), \theta; h)) \Big|_{y'=y, \theta=d\varphi(y)^{-T}(\eta-d\psi(y))} + \mathcal{O}(h^N)_{S_{L_0, \rho}^{\text{comp}}(\mathbb{R}^n)}.$$

where  $L_j$  are differential operators of order  $2j$  in  $y', \theta$  depending on  $\varphi, \psi$  and  $L_0 = 1$ .

*Proof.* We write

$$\begin{aligned} & B' \text{Op}_h(a) B f(y) \\ &= (2\pi h)^{-n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}((\varphi(y)-\varphi(y'))\cdot\theta+\psi(y)-\psi(y'))} \chi(y)\chi(y') J_\varphi(y') a(\varphi(y), \theta; h) f(y') dy' d\theta \end{aligned}$$

where  $J_\varphi(y') = |\det d\varphi(y')|$ . By oscillatory testing [Zw, Theorem 4.19], we have  $B' \text{Op}_h(a) B = \text{Op}_h(b)$ , where

$$\begin{aligned} & b(y, \eta; h) = e^{-\frac{i}{h}y\cdot\eta} B' \text{Op}_h(a) B (e^{\frac{i}{h}y'\cdot\eta}) \\ &= (2\pi h)^{-n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}((\varphi(y)-\varphi(y'))\cdot\theta+\psi(y)-\psi(y')-(y-y')\cdot\eta)} \chi(y)\chi(y') J_\varphi(y') a(\varphi(y), \theta; h) dy' d\theta, \end{aligned}$$

as long as all derivatives of  $b$  are bounded uniformly on  $\mathbb{R}^{2n}$  for each fixed  $h$ . It then remains to establish the asymptotic expansion for  $b$ , which follows immediately by the method of stationary phase [Zw, Theorem 3.16]. The symbol  $b$  is  $\mathcal{O}(h^\infty)_{\mathcal{S}(\mathbb{R}^{2n})}$  outside of a fixed compact set, therefore it can be cut off to a compactly supported symbol.  $\square$

**Lemma 3.10.** *Assume that  $\varkappa : V \rightarrow V'$ , where  $V, V' \subset T^*\mathbb{R}^n$  are open, is a canonical transformation which preserves the foliation  $L_0$ . Let*

$$B \in I_h^{\text{comp}}(\varkappa), \quad B' \in I_h^{\text{comp}}(\varkappa^{-1}).$$

Take  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ . Then there exists  $b \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  such that

$$\begin{aligned} & B' \text{Op}_h(a) B = \text{Op}_h(b) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}, \\ & b = (a \circ \varkappa) \sigma_h(B' B) + \mathcal{O}(h^{1-\rho})_{S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)}. \end{aligned}$$

Moreover, if  $h_j \rightarrow 0$ ,  $(y_j, \eta_j) \in T^*\mathbb{R}^n$  are some sequences such that  $a$  is  $\mathcal{O}(h^\infty)$  along the sequence  $(\varkappa(y_j, \eta_j), h_j)$  (in the sense of Definition 3.4), then  $b$  is  $\mathcal{O}(h^\infty)$  along the sequence  $(y_j, \eta_j, h_j)$ .

*Proof.* By applying a partition of unity to  $B, B'$  and using pseudolocality of  $\text{Op}_h(a)$  (see (3.7)) and part 2 of Lemma 3.6, we reduce to the case when  $\varkappa$  has the form (3.5) for some  $\varphi : B(y_0, \varepsilon) \rightarrow \mathbb{R}^n$ ,  $\psi \in C^\infty(B(y_0, \varepsilon); \mathbb{R})$ ; we add a constant to  $\psi$  to make sure that the fixed antiderivative on  $\text{Gr}(\varkappa)$  is equal to  $\psi(x)$ . By Lemma 2.1 and the composition property of Fourier integral operators, the products

$$A := e^{\frac{i}{h}\psi} \varphi^* B, \quad A' := B'(\varphi^{-1})^* e^{-\frac{i}{h}\psi},$$

lie in  $\Psi_h^{\text{comp}}(\mathbb{R}^n)$ . (Lemma 2.1 applies since we can insert an element of  $\Psi_h^{\text{comp}}$  in between  $B, B'$  and other factors.) Since  $\text{WF}'_h(B) \subset \text{Gr}(\mathcal{L})$  and  $\text{WF}'_h(B') \subset \text{Gr}(\mathcal{L}')$  are compact, there exists  $\chi \in C_0^\infty(B(y_0, \varepsilon))$  such that

$$B = (\chi \circ \varphi^{-1})B + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}, \quad B' = B'(\chi \circ \varphi^{-1}) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}.$$

Then we write

$$B' \text{Op}_h(a)B = A'(\chi e^{\frac{i}{h}\psi} \varphi^* \text{Op}_h(a)(\varphi^{-1})^* e^{-\frac{i}{h}\psi} \chi)A + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}.$$

By Lemma 3.9, we can write the operator in parentheses on the right-hand side as  $\text{Op}_h(\tilde{a}) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}$  for some  $\tilde{a} \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ ; by Lemma 3.8, we have  $A' \text{Op}_h(\tilde{a})A = \text{Op}_h(b) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}$  for some  $b \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ . The expression for the principal part of  $b$  and the microlocal vanishing statement follow directly from Lemmas 3.8 and 3.9 and the fact that  $\sigma_h(A')\sigma_h(A) = \sigma_h(A'A) = \sigma_h(B'B)$ .  $\square$

**3.3. General calculus.** We now introduce pseudodifferential operators associated to general Lagrangian foliations, starting with the following

**Definition 3.11.** *Let  $M$  be a manifold,  $U \subset T^*M$  an open set,  $L$  a Lagrangian foliation on  $U$ , and  $\rho \in [0, 1)$ . A family of operators*

$$A = A(h) : \mathcal{D}'(M) \rightarrow C_0^\infty(M)$$

*is called a semiclassical pseudodifferential operator with symbol of class  $S_{L, \rho}^{\text{comp}}(U)$  (denoted  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ ) if it can be written in the form*

$$A = \sum_{\ell=1}^N B'_\ell \text{Op}_h(a_\ell) B_\ell + \mathcal{O}(h^\infty)_{\mathcal{D}'(M) \rightarrow C_0^\infty(M)} \quad (3.8)$$

*for some Lagrangian charts  $\mathcal{L}_\ell$ , Fourier integral operators  $B_\ell \in I_h^{\text{comp}}(\mathcal{L}_\ell)$ ,  $B'_\ell \in I_h^{\text{comp}}(\mathcal{L}_\ell^{-1})$ , and symbols  $a_\ell \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ .*

**Lemma 3.12.** *Let  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ . Then:*

1.  *$A$  is bounded on  $L^2$  uniformly in  $h$ , pseudolocal, and  $\text{WF}_h(A) \subset U$  is compact.*
2. *If  $\mathcal{L} : \tilde{U} \rightarrow T^*\mathbb{R}^n$  is a Lagrangian chart and  $B \in I_h^{\text{comp}}(\mathcal{L})$ ,  $B' \in I_h^{\text{comp}}(\mathcal{L}^{-1})$ , then*

$$BAB' = \text{Op}_h(a) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}$$

*for some  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ ,  $\text{supp } a \subset \mathcal{L}(\tilde{U})$ , and for each representation (3.8) of  $A$ ,*

$$a \circ \mathcal{L} = \sigma_h(B'B) \sum_{\ell=1}^N \sigma_h(B'_\ell B_\ell)(a_\ell \circ \mathcal{L}_\ell) + \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}. \quad (3.9)$$

*Moreover, if  $h_j \rightarrow 0$  and  $(x_j, \xi_j) \in U$  are sequences such that for each  $\ell$ , either  $(x_j, \xi_j) \notin \pi_1(\text{WF}'_h(B_\ell)) \cap \pi_2(\text{WF}'_h(B'_\ell))$  for all  $j$  or  $a_\ell \circ \mathcal{L}_\ell$  is  $\mathcal{O}(h^\infty)$  along  $(x_j, \xi_j, h_j)$  for all  $\ell$  in the sense of Definition 3.4, then  $a \circ \mathcal{L}$  is  $\mathcal{O}(h^\infty)$  along  $(x_j, \xi_j, h_j)$  as well.*

*Proof.* 1. This follows immediately from Lemma 3.7 and the properties of  $\text{Op}_h(a)$ ,  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$  established in the paragraph following (3.7).

2. We write  $A$  in the form (3.8), then

$$BAB' = \sum_{j=1}^N (BB'_j) \text{Op}_h(a_\ell)(B_\ell B') + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}.$$

Now, we have  $B_\ell B' \in I_h^{\text{comp}}(\mathcal{K}'_j)$ , where  $\mathcal{K}'_j = \mathcal{K}_j \circ \mathcal{K}^{-1} : \mathcal{K}(\tilde{U} \cap U_\ell) \rightarrow T^*\mathbb{R}^n$  is a symplectomorphism onto its image preserving the foliation  $L_0$  and  $U_\ell$  is the domain of  $\mathcal{K}_j$ . Similarly  $BB'_j \in I_h^{\text{comp}}((\mathcal{K}'_j)^{-1})$ . It remains to apply Lemma 3.10. To see (3.9), we use the following corollary of (2.12):  $\sigma_h(BB'_j B_\ell B') = (\sigma_h(B'_j B_\ell) \sigma_h(B' B)) \circ \mathcal{K}^{-1}$ .  $\square$

We now define the principal symbol and ( $h$ -dependent) microsupport of an operator in  $\Psi_{h, L, \rho}^{\text{comp}}(U)$ :

**Definition 3.13.** Let  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ . We define the principal symbol

$$\sigma_h^L(A) \in S_{L, \rho}^{\text{comp}}(U) / h^{1-\rho} S_{L, \rho}^{\text{comp}}(U)$$

by the following formula valid for any representation (3.8):

$$\sigma_h^L(A) = \sum_{\ell=1}^N \sigma_h(B'_\ell B_\ell)(a_\ell \circ \mathcal{K}_\ell). \quad (3.10)$$

Moreover, if  $h_j \rightarrow 0$  and  $(x_j, \xi_j) \in U$  are some sequences, then we say that  $A = \mathcal{O}(h^\infty)$  microlocally along  $(x_j, \xi_j, h_j)$ , if for each choice of  $\mathcal{K}, B, B'$  in part 2 of Lemma 3.12 and the corresponding symbol  $a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ , the symbol  $a \circ \mathcal{K}$  is  $\mathcal{O}(h^\infty)$  along  $(x_j, \xi_j, h_j)$ .

It follows from Lemma 3.12 that  $\sigma_h^L(A)$  does not depend on the choice of the representation (3.8) of  $A$ .

To construct an operator with given principal symbol and microsupport, we use the quantization map

$$a \in S_{L, \rho}^{\text{comp}}(U) \mapsto \text{Op}_h^L(a) := \sum_{\ell} B'_\ell \text{Op}_h(a_\ell) B_\ell, \quad (3.11)$$

where the sum above has finitely many nonzero terms for each  $a$  and

- $\mathcal{K}_\ell : U_\ell \rightarrow T^*\mathbb{R}^n$  are Lagrangian charts and  $U_\ell$ ,  $\ell \in \mathbb{N}$ , form a locally finite covering of  $U$ ;
- $B_\ell \in I_h^{\text{comp}}(\mathcal{K}_\ell)$ ,  $B'_\ell \in I_h^{\text{comp}}(\mathcal{K}_\ell^{-1})$  are Fourier integral operators such that  $\sigma_h(B'_\ell B_\ell) \in C_0^\infty(U_\ell)$  form a partition of unity:

$$\sum_{\ell} \sigma_h(B'_\ell B_\ell) = 1 \quad \text{on } U;$$

- $a_\ell = (\chi_\ell a) \circ \varkappa_\ell^{-1} \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n)$ , where  $\chi_\ell \in C_0^\infty(U_\ell)$  are some functions equal to 1 near  $\text{supp } \sigma_h(B'_\ell B_\ell)$ .

One can choose  $\varkappa_\ell$  with the required properties by Lemma 3.6; for existence of  $B_\ell, B'_\ell$  see the discussion following (2.13). The quantization map is not canonical as it depends on the choice of  $\varkappa_\ell, B_\ell, B'_\ell, \chi_\ell$ .

Note that if  $A \in \Psi_h^{\text{comp}}(M)$  is a pseudodifferential operator in the standard calculus and  $\text{WF}_h(A) \subset U$ , then  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$  and  $\sigma_h^L(A) = \sigma_h(A)$ . Also, if  $a \in S_h^0(T^*M)$  is a symbol in the standard class supported inside  $U$ , then  $\text{Op}_h^L(a) \in \Psi_h^{\text{comp}}(M)$ . This follows from the composition property of Fourier integral operators together with (2.12).

The basic properties of the symbol map and a quantization map are given by

**Lemma 3.14.** 1. For each  $a \in S_{L, \rho}^{\text{comp}}(U)$ ,

$$\sigma_h^L(\text{Op}_h^L(a)) = a + \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}.$$

2. For each  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ , we have  $\sigma_h^L(A) = \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}$  if and only if  $A \in h^{1-\rho}\Psi_{h, L, \rho}^{\text{comp}}(U)$ .

3. For each  $A \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ , there exists  $a \in S_{L, \rho}^{\text{comp}}(U)$  such that  $A = \text{Op}_h^L(a) + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}$ .

4. For each  $a \in S_{L, \rho}^{\text{comp}}(U)$ , if  $h_j \rightarrow 0$  and  $(x_j, \xi_j) \in U$  are sequences such that  $a$  is  $\mathcal{O}(h^\infty)$  along  $(x_j, \xi_j, h_j)$  (in the sense of Definition 3.4), then  $\text{Op}_h^L(a)$  is  $\mathcal{O}(h^\infty)$  microlocally along  $(x_j, \xi_j, h_j)$  (in the sense of Definition 3.13).

5. Let  $A, B \in \Psi_{h, L, \rho}^{\text{comp}}(U)$ . Then  $AB, A^* \in \Psi_{h, L, \rho}^{\text{comp}}(U)$  and

$$\begin{aligned} \sigma_h^L(AB) &= \sigma_h^L(A)\sigma_h^L(B) + \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}, \\ \sigma_h^L(A^*) &= \overline{\sigma_h^L(A)} + \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}. \end{aligned}$$

*Proof.* 1. This follows immediately from (3.10).

2. Assume that  $\sigma_h^L(A) = \mathcal{O}(h^{1-\rho})_{S_{L, \rho}^{\text{comp}}(U)}$ ; we need to show that  $A \in h^{1-\rho}\Psi_{h, L, \rho}^{\text{comp}}(U)$  (the reverse implication follows directly from (3.10)). Using a pseudodifferential partition of unity, we may assume that  $\text{WF}_h(A)$  is contained in a some open subset  $\tilde{U} \subset U$  such that there exists a Lagrangian chart  $\varkappa : \tilde{U} \rightarrow T^*\mathbb{R}^n$  and Fourier integral operators

$$B \in I_h^{\text{comp}}(\varkappa), \quad B' \in I_h^{\text{comp}}(\varkappa^{-1}); \quad B'B = 1 + \mathcal{O}(h^\infty) \quad \text{microlocally near } \text{WF}_h(A).$$

Then

$$A = B'(BAB')B + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}.$$

However, by part 2 of Lemma 3.12 we have

$$BAB' = \text{Op}_h(a) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}, \quad a \in S_{L_0, \rho}^{\text{comp}}(T^*\mathbb{R}^n),$$

and since  $\sigma_h^L(A) = \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(U)}$ , we have  $a = \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(T^*\mathbb{R}^n)}$ . Therefore,  $a = h^{1-\rho}b$  for some  $b \in S_{L,\rho}^{\text{comp}}(T^*\mathbb{R}^n)$  and

$$A = h^{1-\rho}(B' \text{Op}_h(b)B + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}) \in h^{1-\rho}\Psi_{h,L,\rho}^{\text{comp}}(U).$$

3. Put  $a_0 = \sigma_h^L(A)$ . Then  $A = \text{Op}_h^L(a_0) + \mathcal{O}(h^{1-\rho})_{\Psi_{h,L,\rho}^{\text{comp}}(U)}$ , therefore  $A = \text{Op}_h^L(a_0) + h^{1-\rho}A_1$  for some  $A_1 \in \Psi_{h,L,\rho}^{\text{comp}}(U)$ . By induction we construct a family of operators  $A_j \in \Psi_{h,L,\rho}^{\text{comp}}(U)$ ,  $j \in \mathbb{N}_0$ , such that  $A_0 = A$  and  $A_j = \text{Op}_h^L(\sigma_h^L(A_j)) + h^{1-\rho}A_{j+1}$ . Then we have  $A = \text{Op}_h^L(a) + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}$  where  $a \in S_{h,L,\rho}^{\text{comp}}(U)$  is the following asymptotic sum:

$$a \sim \sum_{j=0}^{\infty} h^{(1-\rho)j} \sigma_h^L(A_j).$$

4,5. These follow from Lemma 3.8 and part 2 of Lemma 3.12.  $\square$

**3.4. Further properties.** We start with an improved bound on the  $L^2$  operator norm:

**Lemma 3.15.** *Let  $A \in \Psi_{h,L,\rho}^{\text{comp}}(U)$ . Then as  $h \rightarrow 0$ ,*

$$\|A\|_{L^2(M) \rightarrow L^2(M)} \leq \sup_U |\sigma_h^L(A)| + o(1).$$

*Proof.* Take  $\varepsilon > 0$  and let  $a := \sigma_h^L(A)$ . It suffices to prove that

$$\limsup_{h \rightarrow 0} \|A\|_{L^2(M) \rightarrow L^2(M)} \leq C_\varepsilon := \sup_U |a| + \varepsilon. \quad (3.12)$$

Define the function  $b$  by

$$b = \sqrt{C_\varepsilon^2 - |a|^2}.$$

Note that  $b = C_\varepsilon$  outside of  $\text{supp } a$  and  $C_\varepsilon - b \in S_{L,\rho}^{\text{comp}}(U)$ . Take the following quantization of  $b$ :

$$B := C_\varepsilon - \text{Op}_h^L(C_\varepsilon - b)$$

and note that  $B$  is bounded on  $L^2$ . Since  $|a|^2 + |b|^2 = C_\varepsilon^2$ , we have

$$A^*A + B^*B = C_\varepsilon^2 + \mathcal{O}(h^{1-\rho})_{L^2 \rightarrow L^2}.$$

By applying this to  $u \in L^2$  and taking the scalar product with  $u$  itself, we get

$$\|Au\|_{L^2}^2 \leq C_\varepsilon^2 \|u\|_{L^2}^2 + \mathcal{O}(h^{1-\rho}) \|u\|_{L^2}^2,$$

which implies (3.12).  $\square$

We next give a version of the elliptic parametrix construction:

**Lemma 3.16.** *Assume that  $A, B \in \Psi_{h,L,\rho}^{\text{comp}}(U)$  and  $B$  is elliptic on the microsupport of  $A$  in the following sense: there exists  $\varepsilon > 0$  such that for each sequences  $h_j \rightarrow 0$  and  $(x_j, \xi_j) \in U$ , if  $|\sigma_h^L(B)(x_j, \xi_j; h_j)| \leq \varepsilon$ , then  $A$  is  $\mathcal{O}(h^\infty)$  microlocally along  $(x_j, \xi_j; h_j)$ . Then there exists*

$$Q \in \Psi_{h,L,\rho}^{\text{comp}}(U), \quad A = QB + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}.$$

*Proof.* Let  $a_0 = \sigma_h^L(A)$ ,  $b_0 = \sigma_h^L(B)$ . We first show that there exists  $q_0 \in S_{L,\rho}^{\text{comp}}(U)$  such that  $a_0 = q_0 b_0 + \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(U)}$ . For that, let  $\chi \in C_0^\infty(-\varepsilon, \varepsilon)$  be equal to 1 near the origin. Then from the ellipticity assumption we have

$$a_0 \chi(|b_0|) = \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(U)}$$

and it remains to put  $q_0 := a_0(1 - \chi(|b_0|))/b_0$ .

Put  $Q_0 = \text{Op}_h^L(q_0)$ , then

$$A = Q_0 B + \mathcal{O}(h^{1-\rho})_{\Psi_{h,L,\rho}^{\text{comp}}(U)}.$$

We write  $A = Q_0 B + h^{1-\rho} R_1$  for some  $R_1 \in \Psi_{h,L,\rho}^{\text{comp}}(U)$ . By part 4 of Lemma 3.14,  $B$  is elliptic on the microsupport of  $Q_0$ ; therefore, it is elliptic on the microsupport of  $R_1$ . Repeating the above process, we find a sequence  $Q_j = \text{Op}_h^L(q_j)$ ,  $R_j \in \Psi_{h,L,\rho}^{\text{comp}}(U)$  such that  $A = R_0$  and

$$R_j = Q_j B + h^{1-\rho} R_{j+1}, \quad j = 0, 1, \dots$$

It remains to take  $Q = \text{Op}_h^L(q)$ , where  $q \in S_{L,\rho}^{\text{comp}}(U)$  is the asymptotic sum

$$q \sim \sum_{j=0}^{\infty} h^{(1-\rho)j} q_j. \quad \square$$

Finally, we give a version of Egorov's theorem for the  $\Psi_{h,L,\rho}^{\text{comp}}$  calculus:

**Lemma 3.17.** *Let  $L$  be a Lagrangian foliation on  $U \subset T^*M$ ,  $P \in \Psi_h^{\text{comp}}(M)$ , the principal symbol  $p = \sigma_h(P)$  be real-valued (and  $h$ -independent), and*

$$L_{(x,\xi)} \subset \ker dp(x, \xi) \quad \text{for each } (x, \xi) \in U. \quad (3.13)$$

*Let  $A \in \Psi_{h,L,\rho}^{\text{comp}}(U)$  and take  $T > 0$  such that  $e^{-tH_p}(\text{WF}_h(A)) \subset U$  for all  $t \in [0, T]$ . Then there exists a family of operators depending smoothly on  $t$*

$$A_t \in \Psi_{h,L,\rho}^{\text{comp}}(U), \quad t \in [0, T], \quad A_0 = A + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty},$$

*such that  $\sigma_h^L(A_t) = \sigma_h^L(A) \circ e^{tH_p} + \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(U)}$  and*

$$ih\partial_t A_t + [P, A_t] = \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}. \quad (3.14)$$

*Moreover, if  $t$  is fixed and  $h_j \rightarrow 0$ ,  $(x_j, \xi_j) \in U$  are sequences such that  $A$  is  $\mathcal{O}(h^\infty)$  microlocally along  $(x_j, \xi_j, h_j)$ , then  $A_t$  is  $\mathcal{O}(h^\infty)$  microlocally along  $(e^{-tH_p}(x_j, \xi_j), h_j)$ .*

*Proof.* First of all, by (3.13) the Hamiltonian vector field  $H_p$  lies in  $L$ . Therefore, the Hamiltonian flow  $e^{tH_p}$  preserves  $L$ , and for each  $a \in S_{L,\rho}^{\text{comp}}(U)$ , both  $H_p a$  and  $a \circ e^{tH_p}$  (as long as  $e^{-tH_p}(\text{supp } a) \subset U$ ) lie in  $S_{L,\rho}^{\text{comp}}(U)$  as well. Next, we have for each  $\tilde{A} \in \Psi_{h,L,\rho}^{\text{comp}}(U)$

$$[P, \tilde{A}] = \frac{h}{i} \text{Op}_h^L(H_p \sigma_h^L(\tilde{A})) + \mathcal{O}(h^{2-\rho})_{\Psi_{h,L,\rho}^{\text{comp}}(U)}. \quad (3.15)$$

Indeed, by a pseudodifferential partition of unity and part 2 of Lemma 3.12, it suffices to prove\* that for each function  $f \in C_0^\infty(\mathbb{R}^n)$ ,

$$f(y) \# \tilde{a} - \tilde{a} \# f(y) = \frac{h}{i} \{f(y), \tilde{a}\} + \mathcal{O}(h^2)_{S_{h,L,\rho}^{\text{comp}}(T^*\mathbb{R}^n)}$$

which follows immediately from Lemma 3.8.

Take  $a \in S_{L,\rho}^{\text{comp}}(U)$  such that  $A = \text{Op}_h^L(a) + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C^\infty}$ . Consider the family of operators

$$A_t^{(0)} := \text{Op}_h^L(a_t^{(0)}), \quad a_t^{(0)} = a \circ e^{tH_p}, \quad t \in [0, T].$$

Then (3.15) implies

$$ih\partial_t A_t^{(0)} + [P, A_t^{(0)}] = h^{2-\rho} R_t^{(1)}, \quad R_t^{(1)} \in \Psi_{h,L,\rho}^{\text{comp}}(U).$$

Next, put for  $t \in [-T, T]$ ,

$$A_t^{(1)} := \text{Op}_h^L(a_t^{(1)}), \quad a_t^{(1)} := \int_0^t i\sigma_h^L(R_s^{(1)}) \circ e^{(t-s)H_p} ds.$$

Then

$$a_0^{(1)} = 0, \quad \partial_t a_t^{(1)} = H_p a_t^{(1)} + i\sigma_h^L(R_t^{(1)})$$

and thus (3.15) implies

$$ih\partial_t A_t^{(1)} + [P, A_t^{(1)}] + hR_t^{(1)} = h^{2-\rho} R_t^{(2)}, \quad R_t^{(2)} \in \Psi_{h,L,\rho}^{\text{comp}}(U).$$

Arguing by induction, we construct operators  $A_t^{(j)} = \text{Op}_h^L(a_t^{(j)})$ ,  $R_t^{(j)} \in \Psi_{h,L,\rho}^{\text{comp}}(U)$ ,  $j \in \mathbb{N}_0$ , such that  $R_t^{(0)} = 0$ ,  $A_0^{(0)} = A + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C^\infty}$ ,  $A_0^{(j)} = 0$  for  $j > 0$ , and

$$ih\partial_t A_t^{(j)} + [P, A_t^{(j)}] + hR_t^{(j)} = h^{2-\rho} R_t^{(j+1)}.$$

It remains to put

$$A_t := \text{Op}_h^L(a_t), \quad a_t \sim \sum_{j=0}^{\infty} h^{(1-\rho)j} a_t^{(j)}. \quad \square$$

---

\*We do not obtain the error  $\mathcal{O}(h^2)$  in (3.15) because right-hand side is in  $h\Psi_{h,L,\rho}^{\text{comp}}(U)$  and thus Lemma 3.12 produces an  $\mathcal{O}(h \cdot h^{1-\rho})$  error.

## 4. HYPERBOLIC MANIFOLDS

In this section, we assume that  $(M, g)$  is an  $n$ -dimensional convex co-compact hyperbolic manifold, that is, a quotient  $M = \Gamma \backslash \mathbb{H}^n$  of the hyperbolic space  $\mathbb{H}^n$  by a convex co-compact (geometrically finite) subgroup  $\Gamma$  of the isometry group  $\text{PSO}(1, n)$  of  $\mathbb{H}^n$ . We refer the reader to [Bo07] for the formal definition and properties of these manifolds in the important special case of dimension  $n = 2$  and to [Pe87] for the case of general dimension.

We will use the calculus of §3 to obtain fine microlocal bounds on the scattering resolvent on  $M$ , and prove Theorem 3 using these bounds.

**4.1. Dynamical properties.** Define the function  $p \in C^\infty(T^*M \setminus 0)$  by

$$p(x, \xi) = |\xi|_g, \quad (x, \xi) \in T^*M \setminus 0, \quad (4.1)$$

and let  $X$  be the Hamiltonian vector field of  $p$ . Then

$$e^{tX} : T^*M \setminus 0 \rightarrow T^*M \setminus 0 \quad (4.2)$$

is the homogeneous rescaling of the geodesic flow. Here homogeneity means that  $[X, \xi \cdot \partial_\xi] = 0$  where  $\xi \cdot \partial_\xi$  is the generator of dilations.

In what follows, we will identify the cotangent bundle  $T^*M$  with the tangent bundle  $TM$  using the metric  $g$ .

**4.1.1. Stable/unstable decomposition.** For  $(x, \xi) \in T^*M \setminus 0$ , we decompose the tangent space at  $(x, \xi)$  as follows:

$$T_{(x, \xi)}(T^*M) = \mathbb{R}X \oplus \mathbb{R}(\xi \cdot \partial_\xi) \oplus E_s(x, \xi) \oplus E_u(x, \xi) \quad (4.3)$$

where  $E_s, E_u$  are the  $n - 1$  dimensional stable and unstable bundles, defined in the case  $|\xi|_g = 1$  for instance in [DFG, (3.14)] (recalling the identification  $T^*M \simeq TM$ ), and in general by requiring that they are homogeneous. Note that  $E_s, E_u$  are the images of the stable/unstable bundles of  $\mathbb{H}^n$  (which are also denoted  $E_s, E_u$ ) under the covering map

$$\pi_\Gamma : T^*\mathbb{H}^n \rightarrow T^*M, \quad (4.4)$$

and they are tangent to the level sets of  $p$ .

The subbundles  $E_s, E_u$  are invariant under the flow  $e^{tX}$ . Moreover, the projection map  $T_{(x, \xi)}T^*M \rightarrow T_xM$  is an isomorphism from  $E_s(x, \xi)$  onto the space  $\{\eta \in T_xM \mid \langle \xi, \eta \rangle = 0\}$ . Therefore, we can canonically pull back the metric  $g_x$  to  $E_s(x, \xi)$ . Same is true for  $E_u(x, \xi)$ , and we have (see for instance [DFG, §3.3])

$$|de^{tX}(x)v|_g = \begin{cases} e^t|v|_g, & v \in E_u(x, \xi); \\ e^{-t}|v|_g, & v \in E_s(x, \xi). \end{cases} \quad (4.5)$$

For each  $(x, \xi) \in T^*M \setminus 0$ , consider the *weak stable/unstable subspaces*

$$L_s(x, \xi) := \mathbb{R}X(x, \xi) \oplus E_s(x, \xi), \quad L_u(x, \xi) := \mathbb{R}X(x, \xi) \oplus E_u(x, \xi). \quad (4.6)$$

Define the maps

$$B_{\pm} : T^*\mathbb{H}^n \setminus 0 \rightarrow \mathbb{S}^{n-1} \quad (4.7)$$

as follows: for  $(x, \xi) \in T^*\mathbb{H}^n$ ,  $B_{\pm}(x, \xi)$  is the limit of the projection to the ball model of  $\mathbb{H}^n$  of the geodesic  $e^{tX}(x, \xi)$  as  $t \rightarrow \pm\infty$  – see for instance [DFG, §3.4]. Then the lifts of  $L_s, L_u$  to  $T^*\mathbb{H}^n \setminus 0$  are given by [DFG, (3.25)]

$$\begin{aligned} \pi_{\Gamma}^* L_s(x, \xi) &= \ker dB_+(x, \xi) \cap \ker dp(x, \xi), \\ \pi_{\Gamma}^* L_u(x, \xi) &= \ker dB_-(x, \xi) \cap \ker dp(x, \xi). \end{aligned} \quad (4.8)$$

**Lemma 4.1.**  *$L_s$  and  $L_u$  are Lagrangian foliations on  $T^*M \setminus 0$  in the sense of Definition 3.1.*

*Proof.* We consider the case of  $L_s$ ; the case of  $L_u$  is handled similarly. Using the covering map  $\pi_{\Gamma}$ , we reduce to the case  $M = \mathbb{H}^n$ . The fact that  $L_s$  is integrable follows immediately from (4.8). Since  $\dim L_s = n$ , it remains to show that  $\omega(Y_1, Y_2) = 0$  for each  $Y_1, Y_2 \in L_s(x, \xi)$ , where  $\omega$  is the symplectic form on  $T^*\mathbb{H}^n$ . When  $Y_1 = X(x, \xi)$ , this is immediate since  $L_s(x, \xi) \subset \ker dp(x, \xi)$ . Therefore, we may assume that  $Y_1, Y_2 \in E_s(x, \xi)$ . Since  $e^{tX}$  is a Hamiltonian flow, it is a symplectomorphism, and we find

$$\omega(Y_1, Y_2) = \omega(de^{tX}(x, \xi)Y_1, de^{tX}(x, \xi)Y_2) \leq C |de^{tX}(x, \xi)Y_1|_g \cdot |de^{tX}(x, \xi)Y_2|_g$$

where the constant  $C$  in the last inequality is independent of  $t$  since the isometry group  $\text{PSO}(1, n)$  acts transitively on  $\mathbb{H}^n$  and the lifted action on  $T^*\mathbb{H}^n \setminus 0$  preserves  $\omega$ ,  $E_s$ , and the induced metric on  $E_s$ . Letting  $t \rightarrow +\infty$  and using (4.5), we get  $\omega(Y_1, Y_2) = 0$  as required.  $\square$

The next lemma states that the result of propagating a compactly supported symbol up to almost twice the Ehrenfest time lies in the anisotropic class  $S_{L, \rho}^{\text{comp}}$  from Definition 3.2, where  $L = L_s$  or  $L = L_u$  depending on the direction of propagation. See Appendix A for the proof.

**Lemma 4.2.** *Let  $\chi_1, \chi_2 \in C_0^\infty(T^*M \setminus 0)$  be independent of  $h$  and fix  $\rho \in [0, 1)$ . Then we have uniformly in  $t \in [0, \rho \log(1/h)]$ ,*

$$\chi_2(\chi_1 \circ e^{tX}) \in S_{L_s, \rho}^{\text{comp}}(T^*M \setminus 0), \quad \chi_2(\chi_1 \circ e^{-tX}) \in S_{L_u, \rho}^{\text{comp}}(T^*M \setminus 0).$$

4.1.2. *Infinity and trapping.* Since  $M$  is a convex co-compact hyperbolic manifold, it is *even asymptotically hyperbolic* in the sense of [Gu, Definition 1.2]; more precisely, it is the interior of a compact manifold with boundary  $\overline{M}$  such that near  $\partial\overline{M}$ ,

$$g = \frac{d\tilde{x}^2 + g_1(\tilde{x}^2, \tilde{y})}{\tilde{x}^2},$$

where  $\tilde{x} \geq 0$  is a boundary defining function and  $(\tilde{x}, \tilde{y}) \in (0, \varepsilon) \times \partial \overline{M}$  are some product coordinates on a collar neighborhood of  $\partial \overline{M}$ .

It is shown for example in [DyGu14a, Lemma 7.1] that there exists a function  $r : M \rightarrow \mathbb{R}$  such that  $\{r \leq R\}$  is compact for all  $R$  and the sets  $\{r \leq R\}$  are strictly convex for all  $R \geq 0$ ; that is, if we restrict  $r$  to any geodesic on  $M$  and denote by dots derivatives with respect to the geodesic parameter, then at each point of the geodesic we have

$$r \geq 0, \dot{r} = 0 \implies \ddot{r} > 0. \quad (4.9)$$

In fact, it suffices to take  $r := \tilde{x}^{-1} - r_1$  for a boundary defining function  $\tilde{x}$  of  $\overline{M}$  and large enough constant  $r_1 > 0$ .

We now define the *incoming/outgoing tails*  $\Gamma_{\pm}$  by

$$\Gamma_{\pm} = \{(x, \xi) \in T^*M \setminus 0 \mid r(e^{tX}(x, \xi)) \text{ is bounded as } t \rightarrow \mp\infty\}. \quad (4.10)$$

Define also the *trapped set*  $K = \Gamma_+ \cap \Gamma_-$ . It follows from (4.9) that  $\Gamma_{\pm}$  are closed subsets of  $T^*M \setminus 0$  and  $K \subset \{r < 0\}$ , see for instance [DyGu14a, §4.1]. We assume that  $K \neq \emptyset$  (in the case when  $K = \emptyset$ ,  $M$  is known to have an arbitrarily large essential spectral gap, see for instance [Va13b, (1.1)]).

Recall that  $M = \Gamma \backslash \mathbb{H}^n$ , where  $\Gamma$  is a convex co-compact group of hyperbolic isometries. Define the *limit set*  $\Lambda_{\Gamma} \subset \mathbb{S}^{n-1}$  as follows: for each  $x \in \mathbb{H}^n$ ,

$$\Lambda_{\Gamma} = \overline{\{\gamma \cdot x \mid \gamma \in \Gamma\}} \cap \mathbb{S}^{n-1} \quad (4.11)$$

where we use the ball model of the hyperbolic space and the closure is taken in the closed ball in  $\mathbb{R}^n$ . The resulting set is closed and independent of the choice of  $x$ ; see for instance [Pa76b] and [Bo07, Lemma 2.8] for the case of  $n = 2$  and [Su] for general  $n$ .

For each  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$ , we have (see Appendix A for the proof)

$$\begin{aligned} \pi_{\Gamma}(x, \xi) \in \Gamma_+ &\iff B_-(x, \xi) \in \Lambda_{\Gamma}, \\ \pi_{\Gamma}(x, \xi) \in \Gamma_- &\iff B_+(x, \xi) \in \Lambda_{\Gamma}, \end{aligned} \quad (4.12)$$

where the maps  $B_{\pm}$  are defined in (4.7) and  $\pi_{\Gamma}$  is defined in (4.4).

The following statement, when combined with (4.12), implies that for a trajectory  $(x(t), \xi(t))$  of  $e^{tX}$  on  $T^*M \setminus 0$  which stays in a fixed compact set for all  $t \in [0, T]$ , the point  $(x(0), \xi(0))$  is  $\mathcal{O}(e^{-T})$  close to  $\Gamma_-$  and the point  $(x(T), \xi(T))$  is  $\mathcal{O}(e^{-T})$  close to  $\Gamma_+$ . See Appendix A for the proof.

**Lemma 4.3.** *Let  $V \subset T^*\mathbb{H}^n \setminus 0$  be a compact set. Then there exists a constant  $C$  such that for each  $t \geq 0$ ,*

$$(x, \xi) \in V, \quad \pi_{\Gamma}(e^{\pm tX}(x, \xi)) \in \pi_{\Gamma}(V) \implies d(B_{\pm}(x, \xi), \Lambda_{\Gamma}) \leq Ce^{-t}.$$

Here  $d(\cdot, \cdot)$  is the Euclidean distance function on  $\mathbb{S}^{n-1}$ .

**4.2. Scattering resolvent.** Consider the Laplace–Beltrami operator  $\Delta$  on  $(M, g)$  and its  $L^2$  resolvent

$$R(\lambda) = \left( -\Delta - \frac{(n-1)^2}{4} - \lambda^2 \right)^{-1} : L^2(M) \rightarrow H^2(M), \quad \text{Im } \lambda > 0,$$

which may have finitely many poles corresponding to eigenvalues of  $-\Delta$  on the interval  $[0, \frac{(n-1)^2}{4})$ . Then  $R(\lambda)$  continues meromorphically with poles of finite rank as a family of operators

$$R(\lambda) : L^2_{\text{comp}}(M) \rightarrow H^2_{\text{loc}}(M), \quad \lambda \in \mathbb{C}. \quad (4.13)$$

A related question of continuation of Eisenstein series was studied by Patterson [Pa75, Pa76a] in dimension 2 and Perry [Pe87, Pe89] in higher dimensions. The continuation of  $R(\lambda)$  was established by Mazzeo–Melrose [MaMe] and Guillarmou [Gu] for general (even) asymptotically hyperbolic manifolds, and by Guillopé–Zworski [GuZw] for manifolds of constant curvature near infinity. We refer the reader to [Bo07, Chapter 6] for the proof in dimension 2 and an overview of the history of the subject.

To study essential spectral gaps, we write

$$\lambda = h^{-1} - i\nu, \quad \nu \in [-1, \beta - \varepsilon], \quad (4.14)$$

where the semiclassical parameter  $h > 0$  needs to be small enough for the argument to work. (Resolvent bounds for negative  $\text{Re } \lambda$  follow from bounds for positive  $\text{Re } \lambda$  since  $R(\lambda)^* = R(-\bar{\lambda})$ .) We introduce the semiclassical resolvent

$$R_h(\omega) := h^{-2}R(\lambda), \quad \omega := h\lambda = 1 - ih\nu.$$

To derive high frequency estimates near infinity, we use the construction of the meromorphic continuation of the resolvent due to Vasy [Va13a, Va13b]. See in particular [Va13b, §5.1] and also [DaDy, Lemma 2.1], [Dy, §4.4]. The book [DyZw, Chapter 5] provides a detailed account of a slightly modified version of Vasy’s method, which could be used in the present paper, and [Zw15] gives a short self-contained introduction in the nonsemiclassical case. (For the constant curvature case considered here, one could alternatively apply complex scaling, see Zworski [Zw99] and Datchev [Da].) Specifically, we write

$$R_h(\omega)f = \psi_1(\mathcal{P}_h(\omega)^{-1}\psi_2f)|_M, \quad f \in C_0^\infty(M).$$

Here  $\psi_j \in C^\infty(M)$ ,  $j = 1, 2$ , are certain nonvanishing functions depending on  $\omega, h$  and  $\mathcal{P}_h(\omega) \in \Psi_h^2(M_{\text{ext}})$  is a certain family of semiclassical pseudodifferential operators on a compact manifold  $M_{\text{ext}}$  containing  $M$  as an open subset; we have

$$(\mathcal{P}_h(\omega)u)|_M = \psi_2 \left( -h^2\Delta - \frac{h^2(n-1)^2}{4} - \omega^2 \right) \psi_1(u|_M), \quad u \in C^\infty(M_{\text{ext}}). \quad (4.15)$$

Moreover,  $\mathcal{P}_h(\omega)$  is a Fredholm operator between the spaces

$$\{u \in H_h^s(M_{\text{ext}}) \mid \mathcal{P}_h(\omega)u \in H_h^{s-1}(M_{\text{ext}})\} \rightarrow H_h^{s-1}(M_{\text{ext}})$$

provided that  $s > 0$  is large enough depending on  $\beta$ ; the inverse  $\mathcal{P}_h(\omega)^{-1} : H_h^{s-1}(M_{\text{ext}}) \rightarrow H_h^s(M_{\text{ext}})$  is meromorphic in  $\omega$  with poles of finite rank (if we treat  $\omega$  and  $h$  as independent parameters).

For each fixed  $r_0 > 0$  we may arrange so that  $\psi_1 = \psi_2 = 1$  on  $\{r \leq r_0\}$ , see for instance the paragraph preceding [Va13b, (3.14)]. Therefore, to show the resolvent bound (1.3) it suffices to prove the estimate

$$\|u\|_{H_h^s(M_{\text{ext}})} \leq Ch^{-1-2\max(0,\nu)-\varepsilon} \|f\|_{H_h^{s-1}(M_{\text{ext}})} \quad (4.16)$$

when  $h$  is small enough depending on  $\varepsilon$  and

$$\mathcal{P}_h(\omega)u = f, \quad u \in H_h^s(M_{\text{ext}}), \quad f \in H_h^{s-1}(M_{\text{ext}}). \quad (4.17)$$

(The resulting  $H_h^{s-1} \rightarrow H_h^s$  estimate on  $\chi R_h(\omega)\chi$  can be converted to an  $L^2 \rightarrow L^2$  estimate using the elliptic parametrix of  $-h^2\Delta - \frac{h^2(n-1)^2}{4} - \omega^2$  near the fiber infinity, see for instance [Dy, Proposition 3.3].)

We use the following outgoing estimates on the operator  $\mathcal{P}_h(\omega)$ . Their meaning is as follows: since  $R_h(\omega)$  is the *outgoing* resolvent (in the sense that it maps compactly supported functions on  $M$  to functions with outgoing behavior at the infinity of  $M$ ), it should be *semiclassically outgoing*, that is propagate singularities in the forward direction along the geodesic flow. In particular if  $\tilde{u} = R_h(\omega)\tilde{f}$  and  $\tilde{f} = \mathcal{O}(h^\infty)$ , then  $\text{WF}_h(\tilde{u})$  is contained in the outgoing tail  $\Gamma_+$  (as follows from (4.19) below). Moreover, if we control  $u$  near the trapped set then we can bound its norm everywhere (as follows from (4.20) below).

**Lemma 4.4.** *For each  $u, f$  satisfying (4.17), we have the following estimates:*

1. Assume that  $A_1 \in \Psi_h^0(M_{\text{ext}})$ ,  $\text{WF}_h(A_1) \subset \{r \leq r_0\} \subset \overline{T^*M}$ , and

$$\text{WF}_h(A_1) \cap \Gamma_+ \cap \{|\xi|_g = 1\} = \emptyset. \quad (4.18)$$

Then

$$\|A_1 u\|_{H_h^s(M_{\text{ext}})} \leq Ch^{-1} \|f\|_{H_h^{s-1}(M_{\text{ext}})} + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})}. \quad (4.19)$$

2. Assume that  $A_2 \in \Psi_h^{\text{comp}}(M_{\text{ext}})$  is elliptic on  $K \cap \{|\xi|_g = 1\}$ . Then

$$\|u\|_{H_h^s(M_{\text{ext}})} \leq C \|A_2 u\|_{L^2} + Ch^{-1} \|f\|_{H_h^{s-1}(M_{\text{ext}})}. \quad (4.20)$$

**Remark.** The estimates (4.19), (4.20) make it possible to treat the infinite ends of our manifold as a black box; see [Dy, §4] for a more formal treatment. In particular, our results would apply to any manifold with the same trapping structure as a convex co-compact hyperbolic quotient and infinite ends which satisfy (4.19), (4.20); this includes Euclidean ends [Dy, §4.3] and general even asymptotically hyperbolic ends [Dy, §4.4].

*Sketch of proof.* Both of these statements follow from the elliptic estimate [Dy, Proposition 3.2], propagation of singularities [Dy, Proposition 3.4], and radial points estimates [Va13a, Propositions 2.10 and 2.11] applied to the dynamical picture of the Hamiltonian flow of the principal symbol of  $\mathcal{P}_h(\omega)$  as studied in [Va13a, Va13b]. More precisely, condition (4.18) guarantees that each point in  $\text{WF}_h(A_1)$  either lies in the elliptic set of  $\mathcal{P}_h(\omega)$  or the corresponding backwards Hamiltonian flow line converges to the radial sets, near which  $u$  is controlled when  $s$  is large enough depending on  $\beta$ ; this yields (4.19). Next, each backwards Hamiltonian flow line of  $\mathcal{P}_h(\omega)$  either passes through its elliptic set, or converges to the radial sets, or passes through the elliptic set of  $A_2$ ; this yields (4.20).

The proof (in a modified setting using domains with boundary, which however works equally well for our purposes) is described in detail in [DyZw, §6.2.3]. We also refer the reader to [DaDy, Lemma 4.4] and [Dy, Lemma 4.1] for more details on the dynamics of the flow and to [Dy, Lemmas 4.4 and 4.6] for slightly different proofs involving a semiclassically outgoing parametrix for the resolvent.  $\square$

Finally, we write a pseudodifferential equation (see (4.23) below) which is a direct consequence of (4.17) but more convenient for Lemma 4.5 below because the principal symbol of the associated operator is the function  $p$  given by (4.1). Consider the set

$$W_0 := \{r \leq r_0, |\xi|_g \in [1/2, 2]\} \subset T^*M. \quad (4.21)$$

Take  $P \in \Psi_h^{\text{comp}}(M)$  such that  $P^* = P$  and

$$\begin{aligned} P^2 &= -h^2\Delta + \frac{h^2(n-1)^2}{4} + \mathcal{O}(h^\infty) \quad \text{microlocally near } W_0, \\ \sigma_h(P)(x, \xi) &= p(x, \xi) = |\xi|_g \quad \text{near } W_0. \end{aligned} \quad (4.22)$$

We can construct such an operator following [GrSj, Lemma 4.6]: first take  $P_0 \in \Psi_h^{\text{comp}}(M)$  such that  $P_0^* = P_0$  and  $\sigma_h(P_0) = p$  near  $W_0$ . Denote  $\mathbf{P} := -h^2\Delta + \frac{h^2(n-1)^2}{4}$ , then  $\sigma_h(P_0^2) = \sigma_h(\mathbf{P})$  near  $W_0$  and thus

$$\mathbf{P} = P_0^2 + hR_0 + \mathcal{O}(h^\infty) \quad \text{microlocally near } W_0$$

for some  $R_0 \in \Psi_h^{\text{comp}}(M)$  and  $R_0^* = R_0$ . We next construct  $P_1 \in \Psi_h^{\text{comp}}(M)$  such that  $P_1^* = P_1$  and

$$\mathbf{P} = (P_0 + hP_1)^2 + h^2R_1 + \mathcal{O}(h^\infty) \quad \text{microlocally near } W_0$$

for some  $R_1 \in \Psi_h^{\text{comp}}(M)$  and  $R_1^* = R_1$ ; to do that, it suffices to put  $\sigma_h(P_1) = \sigma_h(R_0)/2p$  near  $W_0$ . Arguing by induction, we construct a family of operators  $P_j \in \Psi_h^{\text{comp}}(M)$  such that  $P_j^* = P_j$  and

$$\mathbf{P} = (P_0 + hP_1 + \cdots + h^m P_m)^2 + \mathcal{O}(h^{m+1}) \quad \text{microlocally near } W_0;$$

it remains to take as  $P$  the asymptotic sum  $P \sim \sum_{j=0}^{\infty} h^j P_j$ .

By (4.15), and since  $\psi_1 = \psi_2 = 1$  near  $\{r \leq r_0\}$ , we have

$$\|B(P^2 - \omega^2)u\|_{L^2} \leq C\|f\|_{H_h^{s-1}(M_{\text{ext}})} + \mathcal{O}(h^\infty)\|u\|_{H_h^s(M_{\text{ext}})}$$

for  $u, f$  satisfying (4.17) and each  $B \in \Psi_h^{\text{comp}}(M)$  such that  $\text{WF}_h(B)$  is contained in some small neighborhood of  $W_0$ . We write  $(P^2 - \omega^2) = (P + \omega)(P - \omega)$  and note that  $P + \omega$  is elliptic on  $W_0$ ; therefore, the elliptic estimate [Dy, Proposition 3.2] gives

$$\|A(P - \omega)u\|_{L^2} \leq C\|f\|_{H_h^{s-1}(M_{\text{ext}})} + \mathcal{O}(h^\infty)\|u\|_{H_h^s(M_{\text{ext}})} \quad (4.23)$$

for each  $A \in \Psi_h^{\text{comp}}(M)$  such that  $\text{WF}_h(A) \subset W_0$ .

**4.3. Second microlocalization of the resolvent.** We now take the first step towards proving a spectral gap, which is to use the calculus of §3 and the Lagrangian foliations  $L_u, L_s$  of (4.6) to obtain fine microlocal estimates on solutions to (4.17). We start with a general propagation estimate:

**Lemma 4.5.** *Let  $a, b \in S_{L,\rho}^{\text{comp}}(T^*M \setminus 0)$  where  $L \in \{L_u, L_s\}$ ,  $\rho \in [0, 1)$ , and fix  $T > 0$ . Assume that  $|a| \leq 1$  everywhere and*

$$e^{-tX}(\text{supp } a) \subset \{b = 1\}; \quad e^{-tX}(\text{supp } a) \subset W_0, \quad t \in [0, T],$$

where  $W_0 \subset T^*M \setminus 0$  is defined in (4.21). Then for each  $\varepsilon_0 > 0$  and each  $u, f$  satisfying (4.17) we have

$$\|\text{Op}_h^L(a)u\|_{L^2} \leq (e^{\nu T} + \varepsilon_0)\|\text{Op}_h^L(b)u\|_{L^2} + Ch^{-1}\|f\|_{H_h^s(M_{\text{ext}})} + \mathcal{O}(h^\infty)\|u\|_{H_h^s(M_{\text{ext}})}.$$

where  $\nu$  is defined in (4.14) and  $\text{Op}_h^L$  is a quantization procedure described in (3.11).

*Proof.* Let  $P \in \Psi_h^{\text{comp}}(M)$  be the operator defined in (4.22). Consider the family of operators  $A_t \in \Psi_{h,L,\rho}^{\text{comp}}(T^*M \setminus 0)$ ,  $t \in [0, T]$ , constructed in Lemma 3.17, with  $A_0 = \text{Op}_h^L(a) + \mathcal{O}(h^\infty)$ ; here (3.13) holds since  $\sigma_h(P) = p$  near  $W_0$  and  $L_u, L_s \subset \ker dp$ .

Using (3.14), (4.23), and the fact that  $P^* = P$ , we write

$$\begin{aligned} \partial_t \|A_t u\|_{L^2}^2 &= 2 \text{Re} \langle \partial_t A_t u, A_t u \rangle \\ &= -\frac{2}{h} \text{Im} \langle [P, A_t] u, A_t u \rangle + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})} \cdot \|A_t u\|_{L^2} \\ &= \frac{2}{h} \text{Im} \langle A_t P u, A_t u \rangle + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})} \cdot \|A_t u\|_{L^2} \\ &= -2\nu \|A_t u\|_{L^2}^2 + (Ch^{-1} \|f\|_{H_h^{s-1}(M_{\text{ext}})} + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})}) \|A_t u\|_{L^2}. \end{aligned}$$

Integrating this, we get

$$\|\text{Op}_h^L(a)u\|_{L^2} \leq e^{\nu T} \|A_T u\|_{L^2} + Ch^{-1} \|f\|_{H_h^{s-1}(M_{\text{ext}})} + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})}. \quad (4.24)$$

Now, it follows from part 4 of Lemma 3.14 and Lemma 3.17 that for each sequences  $h_j \rightarrow 0$  and  $(x_j, \xi_j) \in T^*M \setminus 0$  such that  $e^{TX}(x_j, \xi_j) \notin \text{supp } a(\bullet; h_j)$ , the operator  $A_T$

is  $\mathcal{O}(h^\infty)$  microlocally along  $(x_j, \xi_j, h_j)$  in the sense of Definition 3.13. We then apply Lemma 3.16 to write

$$A_T = Q \text{Op}_h^L(b) + \mathcal{O}(h^\infty)_{\mathcal{D}' \rightarrow C_0^\infty}, \quad Q \in \Psi_{h,L,\rho}^{\text{comp}}(T^*M \setminus 0).$$

Moreover, Lemma 3.17 and the proof of Lemma 3.16 give

$$\sigma_h^L(Q) = (a \circ e^{TX})/b + \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(T^*M \setminus 0)} = a \circ e^{TX} + \mathcal{O}(h^{1-\rho})_{S_{L,\rho}^{\text{comp}}(T^*M \setminus 0)}.$$

By Lemma 3.15, we have  $\|Q\|_{L^2 \rightarrow L^2} \leq 1 + \varepsilon_1$  for each  $\varepsilon_1 > 0$  and  $h$  small enough depending on  $\varepsilon_1$ . Therefore,

$$e^{\nu T} \|A_T u\|_{L^2} \leq (e^{\nu T} + \varepsilon_0) \|\text{Op}_h^L(b)u\|_{L^2} + \mathcal{O}(h^\infty) \|u\|_{H_h^s(M_{\text{ext}})}$$

which together with (4.24) finishes the proof.  $\square$

We can now prove second microlocal estimates on solutions to (4.17). Roughly speaking, in the case  $f = 0$  and  $t = \rho \log(1/h)$  the estimate (4.25) below states that  $u$  is concentrated  $h^\rho$  close to  $\Gamma_+$  (for each  $\rho < 1$ ) and the estimate (4.26) states that the  $u$  has to be of size at least  $h^{\nu\rho+}$  in an  $h^\rho$  neighborhood of  $\Gamma_-$  – see (1.15) and (1.16). In §4.4, we will see that the combination of these two facts with the fractal uncertainty principle implies that  $\nu$  cannot be too small, giving an essential spectral gap.

**Lemma 4.6.** *Let  $\chi \in C_0^\infty(T^*M \setminus 0; [0, 1])$  be equal to 1 near  $K \cap \{|\xi|_g = 1\}$ . Fix  $\rho \in [0, 1)$ . Then there exists  $T > 0$  such that we have for each  $\varepsilon_0 > 0$ , uniformly in  $t \in [T, \rho \log(1/h)]$ , and  $u, f$  satisfying (4.17)*

$$\|\text{Op}_h^{L^u}(\chi(1 - \chi \circ e^{-tX}))u\|_{L^2} \leq Ch^{-1}e^{(\max(0,\nu)+\varepsilon_0)t} \|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty) \|u\|_{H_h^s}, \quad (4.25)$$

$$\|u\|_{H_h^s} \leq Ce^{(\nu+\varepsilon_0)t} \|\text{Op}_h^{L^s}(\chi(\chi \circ e^{tX}))u\|_{L^2} + Ch^{-1}e^{(\max(0,\nu)+\varepsilon_0)t} \|f\|_{H_h^{s-1}}. \quad (4.26)$$

Here  $\chi(1 - \chi \circ e^{-tX}) \in S_{L^u, \rho}^{\text{comp}}(T^*M \setminus 0)$ ,  $\chi(\chi \circ e^{tX}) \in S_{L^s, \rho}^{\text{comp}}(T^*M \setminus 0)$  by Lemma 4.2.

*Proof.* Denote

$$F_+(t) := \|\text{Op}_h^{L^u}(\chi(1 - \chi \circ e^{-tX}))u\|_{L^2},$$

$$F_-(t) := \|\text{Op}_h^{L^s}(\chi(\chi \circ e^{tX}))u\|_{L^2},$$

then it suffices to show that for each  $\varepsilon_0 > 0$  there exists  $T > 0$  such that for all  $t_0 \in [T/2, T]$  and  $t \in [0, \rho \log(1/h)]$ , we have (with constants uniform in  $t_0, t$ )

$$F_+(t+t_0) \leq e^{(\nu+\varepsilon_0)t_0} F_+(t) + Ch^{-1} \|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty) \|u\|_{H_h^s}, \quad (4.27)$$

$$F_-(t) \leq e^{(\nu+\varepsilon_0)t_0} F_-(t+t_0) + Ch^{-1} \|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty) \|u\|_{H_h^s}. \quad (4.28)$$

Indeed, iterating these estimates we get for all  $t \in [T, \rho \log(1/h)]$

$$F_+(t) \leq e^{(\nu+\varepsilon_0)t} F_+(T/2) + Ch^{-1} e^{(\max(0,\nu)+\varepsilon_0)t} \|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty) \|u\|_{H_h^s}, \quad (4.29)$$

$$F_-(0) \leq e^{(\nu+\varepsilon_0)t} F_-(t) + Ch^{-1} e^{(\max(0,\nu)+\varepsilon_0)t} \|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty) \|u\|_{H_h^s}. \quad (4.30)$$

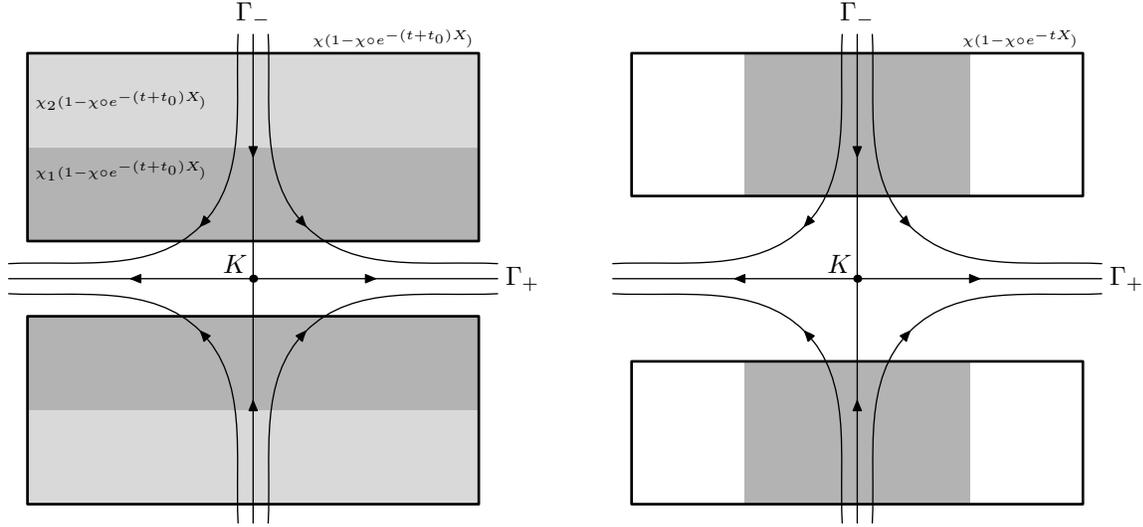


FIGURE 4. An illustration of the proof of (4.27). The function  $\chi(1 - \chi \circ e^{-(t+t_0)X})$  is split into two parts. The part corresponding to  $\chi_2$  is estimated by (4.19) and the darker shaded part corresponding to  $\chi_1$  is transported backwards by the flow to the right half of the figure, where it is covered by  $\chi(1 - \chi \circ e^{-tX})$ .

By (4.31) below, the wavefront set of  $\text{Op}_h^{L^u}(\chi(1 - \chi \circ e^{-TX/2})) \in \Psi_h^{\text{comp}}(M)$  does not intersect  $\Gamma_+ \cap \{|\xi|_g = 1\}$ . By (4.19) (where  $r_0$  is chosen large enough depending on  $\chi$ ) we see that

$$F_+(T/2) \leq Ch^{-1}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}$$

and (4.25) follows from here and (4.29).

Next,  $\text{Op}_h^{L^s}(\chi^2) \in \Psi_h^{\text{comp}}(M)$  is elliptic on  $K \cap \{|\xi|_g = 1\}$ . By (4.20) we get

$$\|u\|_{H_h^s} \leq CF_-(0) + Ch^{-1}\|f\|_{H_h^{s-1}}$$

and (4.26) follows from here and (4.30).

We now prove (4.27). We put  $T := NT_0$ , where  $N$  is a large constant to be chosen later and for each  $(x, \xi) \in \{|\xi|_g = 1\}$  and each  $t, t_1, t_2 \geq T_0 > 0$  we have

$$(x, \xi) \in \Gamma_+ \cap \text{supp } \chi \implies e^{-tX}(x, \xi) \notin \text{supp}(1 - \chi), \quad (4.31)$$

$$(x, \xi) \in e^{t_1X}(\text{supp } \chi) \cap e^{-t_2X}(\text{supp } \chi) \implies (x, \xi) \notin \text{supp}(1 - \chi). \quad (4.32)$$

The existence of such  $T_0$  follows from [DyGu14b, Lemmas 2.3 and 2.4] and the fact that  $\chi = 1$  near  $K \cap \{|\xi|_g = 1\}$ .

We write  $\chi = \chi_1 + \chi_2$  where  $\chi_j \in C_0^\infty(T^*M \setminus 0; [0, 1])$ ,  $\text{supp } \chi_2 \cap \Gamma_+ \cap \{|\xi|_g = 1\} = \emptyset$ , and for each  $t \in [T_0, T + 3T_0]$ ,  $t_1, t_2 \geq T_0$ , and  $(x, \xi) \in T^*M \setminus 0$

$$(x, \xi) \in \text{supp } \chi_1 \implies e^{-tX}(x, \xi) \notin \text{supp}(1 - \chi), \quad (4.33)$$

$$(x, \xi) \in e^{t_1 X}(\text{supp } \chi) \cap e^{-t_2 X}(\text{supp } \chi_1) \implies (x, \xi) \notin \text{supp}(1 - \chi), \quad (4.34)$$

$$(x, \xi) \in e^{t_1 X}(\text{supp } \chi_1) \cap e^{-t_2 X}(\text{supp } \chi) \implies (x, \xi) \notin \text{supp}(1 - \chi). \quad (4.35)$$

Note that (4.34) and (4.35) follow immediately from (4.32) as long as  $\text{supp } \chi_1 \subset \text{supp } \chi$ .

Take  $\chi'_2 \in C_0^\infty(T^*M)$  such that  $\chi'_2 = 1$  near  $\text{supp } \chi_2$  and  $\text{supp } \chi'_2 \cap \Gamma_+ \cap \{|\xi|_g = 1\} = \emptyset$ . By Lemma 3.16 and (4.19) we have

$$\begin{aligned} \|\text{Op}_h^{L_u}(\chi_2(1 - \chi \circ e^{-(t+t_0)X}))u\|_{L^2} &\leq C \|\text{Op}_h^{L_u}(\chi'_2)u\|_{L^2} + \mathcal{O}(h^\infty)\|u\|_{H_h^s} \\ &\leq Ch^{-1}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}. \end{aligned} \quad (4.36)$$

Next, we have (see Figure 4)

$$e^{-(t_0+T_0)X}(\text{supp}(\chi_1(1 - \chi \circ e^{-(t+t_0)X}))) \subset \{\chi(1 - \chi \circ e^{-tX}) = 1\}. \quad (4.37)$$

Indeed, let  $(x, \xi) \in \text{supp}(\chi_1(1 - \chi \circ e^{-(t+t_0)X}))$ . Since  $t_0 + T_0 \in [T_0, T + T_0]$ , by (4.33) we have  $\chi(e^{-(t_0+T_0)X}(x, \xi)) = 1$ . It remains to show that  $\chi(e^{-(t+t_0+T_0)X}(x, \xi)) = 0$ . This follows from (4.34) applied to  $e^{-(t+t_0)X}(x, \xi) \in \text{supp}(1 - \chi)$ ,  $t_1 = T_0$ ,  $t_2 = t + t_0$ .

We now apply Lemma 4.5 to (4.37) (where we choose  $r_0$  large enough depending on  $\chi$  and  $T$  and make  $\text{supp } \chi_1 \subset \{|\xi|_g \in [1/2, 2]\}$ ) and get for each fixed  $\varepsilon_1 > 0$ ,

$$\|\text{Op}_h^{L_u}(\chi_1(1 - \chi \circ e^{-(t+t_0)X}))u\|_{L^2} \leq (e^{\nu(t_0+T_0)} + \varepsilon_1)F_+(t) + Ch^{-1}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}.$$

Together with (4.36) this implies (4.27) as long as we have

$$e^{\nu(t_0+T_0)} + \varepsilon_1 \leq e^{(\nu+\varepsilon_0)t_0}. \quad (4.38)$$

By choosing  $\varepsilon_1$  small enough, this reduces to  $\nu T_0 < \varepsilon_0 t_0$ , which follows from the fact that  $t_0 \geq T/2 = NT_0/2$  if we choose  $N$  large enough depending on  $\varepsilon_0, \beta$ .

To show (4.28), we first note that similarly to (4.36),

$$\|\text{Op}_h^{L_s}(\chi_2(\chi \circ e^{tX}))u\|_{L^2} \leq Ch^{-1}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}. \quad (4.39)$$

Next, there exists  $T_1 \in [T_0, 3T_0]$  such that (see Figure 5)

$$e^{-(t_0+T_1)X}(\text{supp}(\chi_1(\chi \circ e^{tX}))) \subset \{\chi(\chi \circ e^{(t+t_0)X}) = 1\}. \quad (4.40)$$

Indeed, let  $(x, \xi) \in \text{supp}(\chi_1(\chi \circ e^{tX}))$ . By (4.33), we have  $\chi(e^{-(t_0+T_1)X}(x, \xi)) = 1$ . It remains to show that  $\chi(e^{(t-T_1)X}(x, \xi)) = 1$ . If  $t \leq 2T_0$ , then we put  $T_1 := t + T_0$  and use (4.33). If  $t \geq 2T_0$ , then we put  $T_1 := T_0$  and apply (4.35) to  $e^{(t-T_0)X}(x, \xi)$ ,  $t_1 = t - T_0$ ,  $t_2 = T_0$ .

Applying Lemma 4.5 to (4.40) we get for each fixed  $\varepsilon_1 > 0$ ,

$$\|\text{Op}_h^{L_s}(\chi_1(\chi \circ e^{tX}))\|_{L^2} \leq (e^{\nu(t_0+T_1)} + \varepsilon_1)F_-(t + t_0) + Ch^{-1}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}.$$

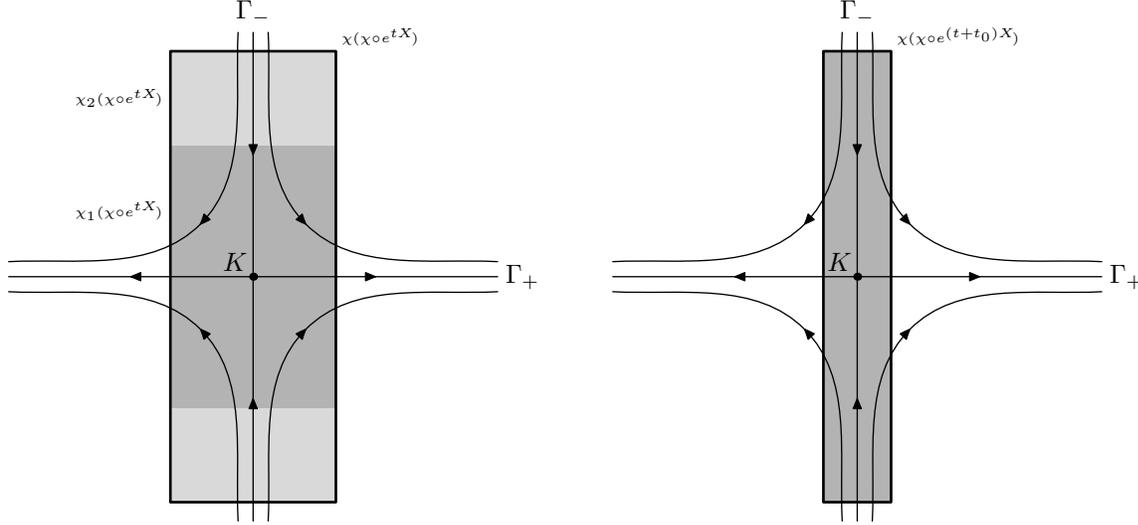


FIGURE 5. An illustration of the proof of (4.28). The function  $\chi(\chi \circ e^{tX})$  is split into two parts. The part corresponding to  $\chi_2$  is estimated by (4.19) and the darker shaded part corresponding to  $\chi_1$  is transported backwards by the flow to the right half of the figure, where it is covered by  $\chi(\chi \circ e^{(t+t_0)X})$ .

Together with (4.39) this implies (4.28) as long as we have

$$e^{\nu(t_0+T_1)} + \varepsilon_1 \leq e^{(\nu+\varepsilon_0)t_0}$$

which is achieved by taking  $N$  large enough similarly to (4.38).  $\square$

**4.4. Reduction to a fractal uncertainty principle.** In this section, we prove Theorem 3. We start by constructing symplectomorphisms

$$\mathcal{K}^\pm : T^*\mathbb{H}^n \setminus 0 \rightarrow T^*(\mathbb{R}_w^+ \times \mathbb{S}_y^{n-1}) \quad (4.41)$$

which map the weak stable/unstable Lagrangian foliations  $L_s, L_u$  defined in (4.6) to the vertical foliation on  $T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ :

$$(\mathcal{K}^+)_*L_u = (\mathcal{K}^-)_*L_s = L_V := \ker(dw) \cap \ker(dy). \quad (4.42)$$

Recall the symbol  $p : T^*\mathbb{H}^n \setminus 0 \rightarrow (0, \infty)$  and the maps  $B_\pm : T^*\mathbb{H}^n \setminus 0 \rightarrow \mathbb{S}^{n-1}$  defined in (4.1) and (4.7). For  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$ , put

$$G_\pm(x, \xi) = p(x, \xi)\mathcal{G}(B_\pm(x, \xi), B_\mp(x, \xi)) \in T_{B_\pm(x, \xi)}^*\mathbb{S}^{n-1}$$

where  $\mathcal{G}$  is defined in (1.19). See Figure 6.

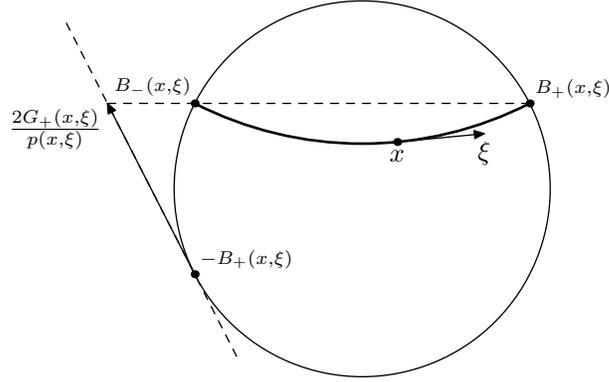


FIGURE 6. The points  $B_{\pm}(x, \xi)$  and the vector  $G_+(x, \xi)$  in the ball model of the hyperbolic space.

Denote by  $\mathcal{P}(x, y)$  the (two-dimensional version of) Poisson kernel, defined on the ball model of  $\mathbb{H}^n$  by

$$\mathcal{P}(x, y) = \frac{1 - |x|^2}{|x - y|^2}, \quad x \in \mathbb{H}^n, \quad y \in \mathbb{S}^{n-1}. \quad (4.43)$$

The symplectomorphisms  $\varkappa^{\pm}$  are constructed in the following lemma; see Appendix A for the proof. Note that (4.42) follows immediately from (4.44) and (4.8).

**Lemma 4.7.** *The maps*

$$\varkappa^{\pm} : (x, \xi) \mapsto (p(x, \xi), B_{\mp}(x, \xi), \pm \log \mathcal{P}(x, B_{\mp}(x, \xi)), \pm G_{\mp}(x, \xi)) \quad (4.44)$$

are exact symplectomorphisms from  $T^*\mathbb{H}^n \setminus 0$  onto  $T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ .

**Remark.** The coordinates  $\varkappa^{\pm}(x, \xi) = (w, y, \theta, \eta)$  can be interpreted as follows:

- $y, \eta$  determine the geodesic  $\gamma(t) = e^{tX}(x, \xi)$  up to shifting  $t$  and rescaling  $\xi$ , in particular  $y$  gives the limit of the geodesic  $\gamma(t)$  as  $t \rightarrow \pm\infty$ ;
- $w$  is the length of  $\xi$ , corresponding to the energy of the geodesic  $\gamma(t)$ ;
- $\theta$  satisfies  $\theta(\gamma(t)) = \theta(\gamma(0)) - t$  and thus determines the position of  $(x, \xi)$  on the geodesic  $\gamma(t)$ .

We next consider the symplectomorphism

$$\widehat{\varkappa} := \varkappa^+ \circ (\varkappa^-)^{-1} : T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1}) \rightarrow T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1}). \quad (4.45)$$

The next lemma, proved in Appendix A, constructs a generating function for  $\widehat{\varkappa}$ :

**Lemma 4.8.** *Consider the following function on  $\mathbb{R}_w^+ \times \mathbb{S}_{\Delta}^{n-1}$ :*

$$\Theta(w, y, y') = w \log \frac{|y - y'|^2}{4} \quad (4.46)$$

where  $|y - y'|$  denotes Euclidean distance on  $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ . Then for each  $(w, y, \theta, \eta)$  and  $(w, y', \theta', \eta')$  in  $T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ , the following two statements are equivalent:

$$(w, y', \theta', \eta') = \widehat{\varkappa}(w, y, \theta, \eta); \quad (4.47)$$

$$\theta - \theta' = \partial_w \Theta(w, y, y'), \quad \eta = \partial_y \Theta(w, y, y'), \quad \eta' = -\partial_{y'} \Theta(w, y, y'). \quad (4.48)$$

Moreover, the antiderivative for  $\widehat{\varkappa}$  defined as the sum of antiderivatives for  $\varkappa^+$  and  $(\varkappa^-)^{-1}$  (see §2.2) is equal to the pullback of  $\Theta$  to the graph  $\text{Gr}(\widehat{\varkappa})$ .

Using Lemma 4.8 and the theory presented in §2.2, we characterize Fourier integral operators associated to  $\widehat{\varkappa}^{-1}$ :

**Lemma 4.9.** *Assume that  $B \in I_h^{\text{comp}}(\widehat{\varkappa}^{-1})$ . Then we have*

$$B = A\widetilde{\mathcal{B}}_\chi + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$$

for some  $A \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ ,  $\chi \in C_0^\infty(\mathbb{S}_\Delta^{n-1})$ , and

$$\widetilde{\mathcal{B}}_\chi v(w, y) = (2\pi h)^{\frac{1-n}{2}} \int_{\mathbb{S}^{n-1}} \left| \frac{y - y'}{2} \right|^{2iw/h} \chi(y, y') v(w, y') dy'$$

where  $\mathbb{S}_\Delta^{n-1}$  is defined in (1.7),  $|y - y'|$  denotes the Euclidean distance, and  $dy'$  is the standard volume form on the sphere.

*Proof.* For the function  $\Theta$  defined in (4.46), we have

$$\widetilde{\mathcal{B}}_\chi v(w, y) = (2\pi h)^{\frac{1-n}{2}} \int_{\mathbb{S}^{n-1}} e^{\frac{i}{h}\Theta(w, y, y')} \chi(y, y') v(w, y') dy'.$$

For  $A \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ , the operator  $A\widetilde{\mathcal{B}}_\chi$  is given by the following formula modulo an  $\mathcal{O}(h^\infty)_{\Psi^{-\infty}}$  remainder:

$$A\widetilde{\mathcal{B}}_\chi v(w, y) = (2\pi h)^{-\frac{n+1}{2}} \int_{\mathbb{S}^{n-1} \times T^*\mathbb{R}^+} e^{\frac{i}{h}((w-w')\theta + \Theta(w', y, y'))} b(w, \theta, y, y'; h) v(w', y') dy' dw' d\theta$$

where  $b$  is a compactly supported symbol on  $\mathbb{R}_w^+ \times \mathbb{R}_\theta \times \mathbb{S}_\Delta^{n-1}$  such that

$$b(w, \theta, y, y'; 0) = \sigma_h(A)(w, y, \theta, \partial_y \Theta(w, y, y')) \chi(y, y').$$

To see this, it suffices to choose some local coordinates on  $\mathbb{S}^{n-1}$ , take  $A = \text{Op}_h(a)$  for some compactly supported symbol  $a(w, y, \theta, \eta; h)$ , write

$$\begin{aligned} A\widetilde{\mathcal{B}}_\chi v(w, y) &= (2\pi h)^{\frac{1-3n}{2}} \int e^{\frac{i}{h}((w-w')\theta + (y-y'') \cdot \eta + \Theta(w', y'', y'))} \\ &\quad a(w, y, \theta, \eta; h) \chi(y'', y') v(w', y') dw' dy'' d\theta d\eta dy' \end{aligned}$$

where the integral is taken over  $\mathbb{R}_w^+ \times \mathbb{S}_{y''}^{n-1} \times \mathbb{R}_{\theta, \eta}^n \times \mathbb{S}_{y'}^{n-1}$ , and apply the method of stationary phase in the  $(y'', \eta)$  variables.

Now, let  $B \in I_h^{\text{comp}}(\widehat{\mathfrak{z}}^{-1})$ . Fix  $\chi \in C_0^\infty(\mathbb{S}^{n-1})$  such that

$$\text{WF}_h(B) \subset U_\chi := \varphi_\Theta(\{(w, \theta, y, y') \mid \chi(y, y') \neq 0\}) \quad (4.49)$$

where  $\varphi_\Theta : \mathbb{R}_w^+ \times \mathbb{R}_\theta \times \mathbb{S}_\Delta^{n-1} \rightarrow \text{Gr}(\widehat{\mathfrak{z}}^{-1})$  is the diffeomorphism constructed using (4.48). By Lemma 4.8, the function

$$\Phi : (w, y, w', y', \theta) \in \mathbb{R}_w^+ \times \mathbb{R}_{w'}^+ \times \mathbb{S}_\Delta^{n-1} \times \mathbb{R}_\theta \mapsto (w - w')\theta + \Theta(w', y, y')$$

parametrizes  $\widehat{\mathfrak{z}}^{-1}$  in the sense of (2.7). Therefore, by (2.8) we may write for some compactly supported symbol  $\tilde{b}$  on the domain of  $\Phi$ , modulo  $\mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ ,

$$Bv(w, y) = (2\pi h)^{-\frac{n+1}{2}} \int_{\mathbb{S}^{n-1} \times T^*\mathbb{R}^+} e^{\frac{i}{h}((w-w')\theta + \Theta(w', y, y'))} \tilde{b}(w, y, w', y', \theta; h) v(w', y') dy' dw' d\theta.$$

Moreover, by (4.49) (which implies that  $B$  is a Fourier integral operator associated to a restriction of  $\widehat{\mathfrak{z}}^{-1}$ ) we can take  $\tilde{b}$  supported inside  $\{\chi(y, y') \neq 0\}$ .

Take  $A_0 \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$  such that for all  $(w, y, \theta, \eta, w', y', \theta', \eta') \in \text{Gr}(\widehat{\mathfrak{z}}^{-1})$ ,

$$\sigma_h(A_0)(w, y, \theta, \eta) = \frac{\tilde{b}(w, y, w', y', \theta; 0)}{\chi(y, y')}.$$

Comparing the oscillatory integral expressions for  $A_0 \tilde{\mathcal{B}}_\chi$  and  $B$  and using (2.9), we get

$$B - A_0 \tilde{\mathcal{B}}_\chi \in h I_h^{\text{comp}}(\widehat{\mathfrak{z}}^{-1}).$$

Moreover, we may choose  $A_0$  so that  $\text{WF}_h(B - A_0 \tilde{\mathcal{B}}_\chi) \subset U_\chi$ . Replacing  $B$  with  $h^{-1}(B - A_0 \tilde{\mathcal{B}}_\chi)$  and arguing by induction, we construct  $A_j \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$  such that

$$B - \sum_{j=0}^N h^j A_j \tilde{\mathcal{B}}_\chi \in h^{N+1} I_h^{\text{comp}}(\widehat{\mathfrak{z}}^{-1}).$$

Then  $B = A \tilde{\mathcal{B}}_\chi + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ , where  $A \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$  is defined by the asymptotic sum  $A \sim \sum_{j=0}^\infty h^j A_j$ .  $\square$

We now reformulate the fractal uncertainty principle of Definition 1.1 as follows:

**Lemma 4.10.** *Assume  $\Lambda_\Gamma, \beta, \varepsilon > 0, \rho \in (0, 1)$  are such that (1.8) holds for all  $C_1, \chi$ . With  $L_V$  defined in (4.42), let*

$$A_\pm \in \Psi_{h, L_V, \rho}^{\text{comp}}(T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})), \quad B \in I_h^{\text{comp}}(\widehat{\mathfrak{z}}^{-1})$$

and assume that for some constant  $C_2$ ,  $A_+$  and  $A_-$  are  $\mathcal{O}(h^\infty)$ , in the sense of Definition 3.13, along every sequence  $(w_j, y_j, \theta_j, \eta_j, h_j)$  such that  $d(y_j, \Lambda_\Gamma) > C_2 h_j^\rho$ . Then

$$\|A_- B A_+\|_{L^2 \rightarrow L^2} \leq C h^{\beta - \varepsilon}.$$

*Proof.* We first use Lemma 4.9 to write  $B = A\tilde{\mathcal{B}}_\chi + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}$  for some  $\chi \in C_0^\infty(\mathbb{S}_\Delta^{n-1})$  and  $A \in \Psi_h^{\text{comp}}(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ . The operator  $A_- A \in \Psi_{h, L_V, \rho}^{\text{comp}}$  satisfies the same microlocal vanishing assumption as  $A_-$ , therefore it suffices to show that

$$\|A_- \chi'(w) \tilde{\mathcal{B}}_\chi A_+\|_{L^2 \rightarrow L^2} \leq Ch^{\beta-\varepsilon}. \quad (4.50)$$

Here we may insert some cutoff function  $\chi' \in C_0^\infty(\mathbb{R}^+)$  since  $A$  is compactly supported.

Using Lemma 3.3, take a function  $\chi_0(y; h) \in C^\infty(\mathbb{S}^{n-1})$  such that

$$\text{supp } \chi_0 \subset \Lambda_\Gamma(2C_2 h^\rho), \quad \text{supp}(1 - \chi_0) \cap \Lambda_\Gamma(C_2 h^\rho) = \emptyset, \quad |\partial_y^\alpha \chi_0(y)| \leq C_\alpha h^{-\rho|\alpha|}.$$

Here  $\Lambda_\Gamma(\cdot)$  is defined in (1.4). We claim that

$$A_-(1 - \chi_0(y; h)) = \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}, \quad (1 - \chi_0(y; h))A_+ = \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}. \quad (4.51)$$

Indeed, by a partition of unity we may assume that  $\text{WF}_h(A_-)$  is contained in the cotangent bundle of a coordinate chart on  $\mathbb{R}^+ \times \mathbb{S}^{n-1}$ . By part 2 of Lemma 3.12 (where  $B, B'$  are pullback operators), we can write  $A_- = \text{Op}_h(a_-) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2}$  for some  $a_- \in S_{L_0, \rho}^{\text{comp}}$ . Moreover, by the assumption on  $A_-$ , we see that  $a_-$  is  $\mathcal{O}(h^\infty)$ , in the sense of Definition 3.4, along every sequence  $(w_j, y_j, \theta_j, \eta_j, h_j)$  such that  $y_j \notin \Lambda_\Gamma(C_2 h_j^\rho)$ . Then the first estimate in (4.51) follows from Lemma 3.8 (or rather, its trivial adaptation to the non-compactly supported symbol  $1 - \chi_0(y; h)$ ); the second estimate in (4.51) is proved similarly.

By (4.51), since  $\|A_\pm\|_{L^2 \rightarrow L^2}$  is bounded uniformly in  $h$ , in order to prove (4.50) it suffices to show

$$\|\chi_0(y; h) \chi'(w) \tilde{\mathcal{B}}_\chi \chi_0(y; h)\|_{L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1})} \leq Ch^{\beta-\varepsilon}. \quad (4.52)$$

We calculate

$$\chi_0(y; h) \chi'(w) \tilde{\mathcal{B}}_\chi \chi_0(y; h) v(w, y) = \chi'(w) 4^{-iw/h} \mathcal{B}'_w(v(w, \cdot))(y),$$

where  $\mathcal{B}'_w : L^2(\mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{S}^{n-1})$  is given by

$$\mathcal{B}'_w f(y) = (2\pi h)^{\frac{1-n}{2}} \int_{\mathbb{S}^{n-1}} |y - y'|^{2iw/h} \chi(y, y') \chi_0(y; h) \chi_0(y'; h) f(y') dy'.$$

Replacing  $h$  by  $h/w$  in (1.8) and using that  $\chi_0$  is bounded and supported in  $\Lambda_\Gamma(2C_2 h^\rho)$ , we see that uniformly in  $w \in \text{supp } \chi'$ ,

$$\|\mathcal{B}'_w\|_{L^2(\mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{S}^{n-1})} \leq Ch^{\beta-\varepsilon}$$

and (4.52) follows from here by integration.  $\square$

We are now ready to give

*Proof of Theorem 3.* In order to prove (1.3), it suffices to show the estimate (4.16) for all functions  $u, f$  satisfying (4.17), where (see (4.14))

$$\omega = 1 - ih\nu, \quad \nu \in [-1, \beta - \varepsilon].$$

Take  $\chi_{\pm} \in C_0^\infty(T^*M \setminus 0; [0, 1])$  such that  $\chi_{\pm} = 1$  near  $K \cap \{|\xi|_g = 1\}$ . We also take  $\varepsilon_0 > 0$  small enough depending on  $\varepsilon$  and fix  $\rho \in (0, 1)$  so that (1.8) is satisfied with  $\varepsilon$  replaced by  $\varepsilon_0$ . Put

$$t := \rho \log(1/h), \quad \nu^+ := \max(0, \nu);$$

$$A'_+ := \text{Op}_h^{L_u}(\chi_+(\chi_+ \circ e^{-tX})), \quad A_0 := \text{Op}_h^{L_u}(\chi_+), \quad A'_- := \text{Op}_h^{L_s}(\chi_-(\chi_- \circ e^{tX})).$$

Note that by Lemma 4.2,

$$A'_+ \in \Psi_{h, L_u, \rho}^{\text{comp}}(T^*M \setminus 0), \quad A_0 \in \Psi_h^{\text{comp}}(M), \quad A'_- \in \Psi_{h, L_s, \rho}^{\text{comp}}(T^*M \setminus 0).$$

By Lemma 4.6 we obtain

$$\|(A_0 - A'_+)u\|_{L^2} \leq Ch^{-1-\rho(\nu^++\varepsilon_0)}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s}, \quad (4.53)$$

$$\|u\|_{H_h^s} \leq Ch^{-\rho(\nu^++\varepsilon_0)}\|A'_-u\|_{L^2} + Ch^{-1-\rho(\nu^++\varepsilon_0)}\|f\|_{H_h^{s-1}}. \quad (4.54)$$

We choose  $\chi_{\pm}$  so that  $\chi_+ = 1$  near  $\text{supp } \chi_-$ . Let  $Q \in \Psi_h^{\text{comp}}(M)$  be an elliptic parametrix of  $A_0$  near  $\text{supp } \chi_-$  (see for instance [DyZw, §E.2.2] or [DyZw16, Proposition 2.4]); in particular,

$$QA_0 = 1 + \mathcal{O}(h^\infty) \quad \text{microlocally near } \text{supp } \chi_-.$$

Since  $A'_-$  is pseudolocal and its wavefront set is contained inside  $\text{supp } \chi_-$ , we have

$$\|A'_-(1 - QA_0)u\|_{L^2} = \mathcal{O}(h^\infty)\|u\|_{H_h^s}. \quad (4.55)$$

Take  $\varepsilon_0 < \varepsilon/2$ . We claim that it suffices to prove the bound

$$\|A'_-QA'_+\|_{L^2(M) \rightarrow L^2(M)} \leq Ch^{\beta-\varepsilon_0}. \quad (4.56)$$

Indeed, putting together (4.53)–(4.56) and using that  $\nu \leq \beta - \varepsilon$ , we have

$$\begin{aligned} \|u\|_{H_h^s} &\leq Ch^{-\rho(\nu^++\varepsilon_0)}\|A'_-u\|_{L^2} + Ch^{-1-\rho(\nu^++\varepsilon_0)}\|f\|_{H_h^{s-1}} \\ &\leq Ch^{-\rho(\nu^++\varepsilon_0)}(\|A'_-QA'_+u\|_{L^2} + \|A'_-Q(A_0 - A'_+)u\|_{L^2} \\ &\quad + \|A'_-(1 - QA_0)u\|_{L^2}) + Ch^{-1-\rho(\nu^++\varepsilon_0)}\|f\|_{H_h^{s-1}} \\ &\leq Ch^{\varepsilon-2\varepsilon_0}\|u\|_{H_h^s} + Ch^{-1-2(\nu^++\varepsilon_0)}\|f\|_{H_h^{s-1}} + \mathcal{O}(h^\infty)\|u\|_{H_h^s} \end{aligned}$$

giving (4.16) for  $h$  small enough.

It remains to deduce (4.56) from the fractal uncertainty principle. We may assume that  $\text{WF}_h(Q)$  is contained in a small neighborhood of  $\text{supp } \chi_-$ ; by an appropriate choice of  $\chi_-$ , we may assume that  $\text{WF}_h(Q)$  is contained in a small neighborhood of  $K \cap \{|\xi|_g = 1\}$ . By a partition of unity, it suffices to show that for each  $(x_0, \xi_0) \in K \cap \{|\xi|_g = 1\}$ , there exists a neighborhood  $V$  of  $(x_0, \xi_0)$  such that (4.56) holds for all  $Q \in \Psi_h^{\text{comp}}(M)$  with  $\text{WF}_h(Q) \subset V$ .

Fix  $(x_0, \xi_0) \in K \cap \{|\xi|_g = 1\}$ . Composing the maps  $\varkappa^\pm$  constructed in Lemma 4.7 with a local inverse of the covering map  $\pi_\Gamma$  defined in (4.4), we obtain exact symplectomorphisms

$$\varkappa_0^\pm : U \rightarrow U'_\pm, \quad U \subset T^*M \setminus 0, \quad U'_\pm \subset T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1}), \quad (4.57)$$

for some small neighborhood  $U$  of  $(x_0, \xi_0)$  and some small neighborhoods  $U'_\pm$  of

$$(1, y_0^\pm, \theta_0^\pm, \eta_0^\pm) := \varkappa_0^\pm(x_0, \xi_0).$$

Here  $(w, y)$  are coordinates on  $\mathbb{R}^+ \times \mathbb{S}^{n-1}$  and  $(\theta, \eta)$  are the corresponding dual variables.

Take  $\mathcal{B}_\pm \in I_h^{\text{comp}}(\varkappa_0^\pm)$ ,  $\mathcal{B}'_\pm \in I_h^{\text{comp}}((\varkappa_0^\pm)^{-1})$  quantizing  $\varkappa_0^\pm$  near  $V \times \varkappa_0^\pm(V)$  in the sense of (2.13), where  $V \Subset U$  is a small neighborhood of  $(x_0, \xi_0)$ . Since  $\text{WF}_h(Q) \subset V$  and  $A'_\pm$  are pseudolocal, we have

$$A'_- Q A'_+ = (\mathcal{B}'_- \mathcal{B}_-) A'_- (\mathcal{B}'_- \mathcal{B}_-) (\mathcal{B}'_+ \mathcal{B}_+) Q A'_+ (\mathcal{B}'_+ \mathcal{B}_+) + \mathcal{O}(h^\infty)_{L^2 \rightarrow L^2};$$

therefore, to prove (4.56) it suffices to show that

$$\|A_- B A_+\|_{L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1}) \rightarrow L^2(\mathbb{R}^+ \times \mathbb{S}^{n-1})} \leq Ch^{\beta - \varepsilon_0}, \quad (4.58)$$

where, with  $\widehat{\varkappa}$  defined in (4.45),

$$A_- = \mathcal{B}_- A'_- \mathcal{B}'_-, \quad A_+ = \mathcal{B}_+ Q A'_+ \mathcal{B}'_+, \quad B = \mathcal{B}_- \mathcal{B}'_+ \in I_h^{\text{comp}}(\widehat{\varkappa}^{-1}).$$

By the construction of  $\Psi_{h,L,\rho}^{\text{comp}}$  calculus in §3.3, together with (4.42), we see that

$$A_\pm \in \Psi_{h,L_V,\rho}^{\text{comp}}(T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})).$$

Moreover,  $A_\pm = \mathcal{O}(h^\infty)$  in the sense of Definition 3.13 along each sequence  $(w_j, y_j, \theta_j, \eta_j, h_j)$  such that

$$(w_j, y_j, \theta_j, \eta_j) \notin \varkappa_0^\pm(U \cap e^{\pm\rho \log(1/h_j)X}(\text{supp } \chi_\pm)). \quad (4.59)$$

By Lemma 4.3 and (4.44), we see that (4.59) is satisfied when  $d(y_j, \Lambda_\Gamma) > C_2 h_j^\rho$  for some fixed constant  $C_2$ . Now (4.58) follows from Lemma 4.10.  $\square$

**Remark.** As follows from the proof of Lemma 4.7 in Appendix A, the Fourier integral operators  $\mathcal{B}_\pm$  used in the proof of Theorem 3 are microlocalized versions of the Poisson operators

$$v(w, y) \mapsto u(x) = \int_{\mathbb{R}^+} \int_{\mathbb{S}^{n-1}} \mathcal{P}(x, y)^{\mp iw/h} v(w, y) dy dw.$$

Therefore, conjugation by  $\mathcal{B}_\pm$  is related to the representation of resonant states as images under the Poisson operator of distributions supported on the limit set, see for instance [Bo07, (14.9)] or [BuOl97].

## 5. FRACTAL UNCERTAINTY PRINCIPLE

In this section, we prove Theorem 4. We will not use directly the geometry or the dynamics of the manifold  $M$ , instead relying on the additive structure of the limit set  $\Lambda_\Gamma$  defined in (4.11) and basic harmonic analysis.

**5.1. Basic properties.** We start with some basic facts regarding the fractal uncertainty principle of Definition 1.1. First of all, since  $\mathcal{B}_\chi(h)$  is a semiclassical Fourier integral operator associated to a canonical transformation (see (2.8) and Lemma 4.8), it is bounded on  $L^2$  uniformly in  $h$ . This gives the bound

$$\| \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \mathcal{B}_\chi \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \|_{L^2 \rightarrow L^2} \leq C. \quad (5.1)$$

Combined with Theorem 3, this bound translates to the well-known statement that there are no resonances in the upper half-plane away from the imaginary axis, which is a direct consequence of self-adjointness of the Laplacian on  $L^2(M)$ .

To formulate the next bound, we introduce the parameter

$$\delta \in [0, n - 1] \quad (5.2)$$

defined as the exponent of convergence of Poincaré series: that is,  $\delta$  is the smallest number such that for all  $x, x' \in \mathbb{H}^n$  and  $\operatorname{Re} s > \delta$ , we have

$$\Sigma_p(s; x, x') := \sum_{\gamma \in \Gamma} \exp(-s d_{\mathbb{H}^n}(x, \gamma.x')) < \infty.$$

Here  $d_{\mathbb{H}^n}(\cdot, \cdot)$  stands for the distance function on  $\mathbb{H}^n$  induced by the hyperbolic metric.

The constant  $\delta$  is the Hausdorff dimension of the limit set  $\Lambda_\Gamma$ , see [Bo07, Theorem 14.14] for the case  $n = 2$  and [Su] for general dimensions. It is also the Minkowski dimension of  $\Lambda_\Gamma$ ; in fact, we have the following more precise estimate (which is a form of Ahlfors-David regularity):

$$\begin{aligned} C^{-1} \alpha^{n-1-\delta} (\alpha')^\delta &\leq \mu_L(\Lambda_\Gamma(\alpha) \cap B(y_0, \alpha')) \leq C \alpha^{n-1-\delta} (\alpha')^\delta, \\ 0 < \alpha \leq \alpha' \leq 1, \quad y_0 \in \Lambda_\Gamma, \end{aligned} \quad (5.3)$$

where  $\Lambda_\Gamma(\alpha)$  is defined in (1.4),  $B(y_0, \alpha')$  denotes the ball of radius  $\alpha'$  centered at  $y_0$ ,  $\mu_L$  denotes the Lebesgue measure on  $\mathbb{S}^{n-1}$ , and the constant  $C > 0$  does not depend on  $y_0, \alpha, \alpha'$ . See §7.2 for the proof. By putting  $\alpha' = 1$  in (5.3), we obtain in particular (1.5).

Given (1.5), we may use Schur's lemma [HöIII, Lemma 18.1.12]: the estimate

$$\sup_{y \in \Lambda_\Gamma(C_1 h^\rho)} \int_{\Lambda_\Gamma(C_1 h^\rho)} |\chi(y, y')| dy' \leq C h^{\rho(n-1-\delta)}$$

implies by (1.6) that

$$\| \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \mathcal{B}_\chi \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \|_{L^2 \rightarrow L^2} \leq C h^{\frac{1-n}{2}} h^{\rho(n-1-\delta)}. \quad (5.4)$$

Since we may choose  $\rho$  arbitrarily close to 1, this gives the fractal uncertainty principle with exponent  $\frac{n-1}{2} - \delta$ . By Theorem 3, the bounds (5.1) and (5.4) together translate (with a loss of  $\varepsilon$ ) to the standard spectral gap (1.2).

We finally show that the fractal uncertainty principle cannot hold with  $\beta > \frac{n-1}{2} - \frac{\delta}{2}$ . This threshold corresponds to the Jakobson–Naud conjecture, see §1.1. More precisely, we claim that for each  $\varepsilon > 0$ , there exists  $\chi \in C_0^\infty(\mathbb{S}_\Delta^{n-1})$  and a family  $v(h) \in L^2(\mathbb{S}^{n-1})$  such that for  $h$  small enough,

$$\| \mathbb{1}_{\Lambda_\Gamma(h)} \mathcal{B}_\chi(h) \mathbb{1}_{\Lambda_\Gamma(h)} v(h) \|_{L^2} \geq h^{\frac{n-1}{2} - \frac{\delta}{2} + \varepsilon} \|v(h)\|_{L^2}. \quad (5.5)$$

To prove (5.5), take small  $\tilde{\varepsilon} > 0$ , fix  $y_0, y_1 \in \Lambda_\Gamma$ ,  $y_0 \neq y_1$ , and let  $v(h) = \mathbb{1}_{B(y_0, h^{1+\tilde{\varepsilon}})}$ . Then

$$\|v(h)\|_{L^2} \leq Ch^{(1+\tilde{\varepsilon})(n-1)/2}, \quad \mathbb{1}_{\Lambda_\Gamma(h)} v(h) = v(h). \quad (5.6)$$

Using (1.6), we compute

$$\mathcal{B}_\chi(h)v(y; h) = (2\pi h)^{\frac{1-n}{2}} |y - y_0|^{2i/h} \int_{\mathbb{S}^{n-1}} \left( \frac{|y - y'|}{|y - y_0|} \right)^{2i/h} \chi(y, y') v(y') dy'.$$

We take  $h$ -independent  $\chi$  such that  $\chi(y, y') = 1$  for  $(y, y')$  near  $(y_1, y_0)$ . For  $y$  in a fixed neighborhood of  $y_1$  and  $y' \in \text{supp } v$ , we have  $\left( \frac{|y - y'|}{|y - y_0|} \right)^{2i/h} = 1 + \mathcal{O}(h^\varepsilon)$ , therefore for  $\tilde{\varepsilon}_1$  small enough,

$$|\mathcal{B}_\chi(h)v(y; h)| \geq C^{-1} h^{\frac{1-n}{2}} h^{(1+\tilde{\varepsilon})(n-1)} \quad \text{for } |y - y_1| < \tilde{\varepsilon}_1.$$

By (5.3), we then have

$$\| \mathbb{1}_{\Lambda_\Gamma(h)} \mathcal{B}_\chi(h)v(h) \|_{L^2} \geq C^{-1} h^{\frac{1-n}{2}} h^{(1+\tilde{\varepsilon})(n-1)} h^{\frac{n-1-\delta}{2}},$$

which together with (5.6) implies (5.5) as long as  $\tilde{\varepsilon} < \frac{2}{n-1}\varepsilon$ .

**5.2. Reduction to additive energy.** We now prove Theorem 4. We will take  $\rho$  very close to 1, in particular  $\rho > 1/2$ . Using Lemma 3.3, take  $\psi_0(y; h) \in C^\infty(\mathbb{S}^{n-1}; [0, 1])$  such that

$$\text{supp}(1 - \psi_0) \cap \Lambda_\Gamma(C_1 h^\rho) = \emptyset, \quad \text{supp } \psi_0 \subset \Lambda_\Gamma(h^{\rho/2}); \quad (5.7)$$

$$\sup_{\mathbb{S}^{n-1}} |\partial^\alpha \psi_0| \leq C_\alpha h^{-\frac{\rho}{2}|\alpha|}. \quad (5.8)$$

To show (1.8), it suffices to prove that

$$\| \sqrt{\psi_0} \mathcal{B}_\chi(h) \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \|_{L^2 \rightarrow L^2} \leq Ch^{\beta - \varepsilon},$$

in fact it is enough to prove the following  $T^*T$ -bound:

$$\| \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \mathcal{B}_\chi(h)^* \psi_0 \mathcal{B}_\chi(h) \mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho)} \|_{L^2 \rightarrow L^2} \leq Ch^{2(\beta - \varepsilon)}.$$

By Schur's lemma [HöIII, Lemma 18.1.12] and (1.6) it is enough to prove the Schwartz kernel bound

$$\sup_{y'' \in \Lambda_\Gamma(C_1 h^\rho)} \int_{\Lambda_\Gamma(C_1 h^\rho)} |\mathcal{K}(y, y''; h)| dy \leq Ch^{2(\beta-\varepsilon)} \quad (5.9)$$

where the integral kernel  $\mathcal{K}(y, y''; h)$  of the operator  $\mathcal{B}_\chi(h)^* \psi_0 \mathcal{B}_\chi(h)$  is given by

$$\mathcal{K}(y, y''; h) = h^{1-n} \int_{\mathbb{S}^{n-1}} \left( \frac{|y' - y''|}{|y' - y|} \right)^{2i/h} \chi(y', y'') \overline{\chi(y', y)} \psi_0(y'; h) dy'. \quad (5.10)$$

Morally speaking,  $\mathcal{K}(y, y''; h)$  is the correlation on  $\Lambda_\Gamma(h^{\rho/2})$  between Lagrangian states corresponding to two levels in (1.9) given by  $y_j = y$  and  $y_j = y''$ .

To capture cancellations in the expression (5.9), we use the following precise version of the method of nonstationary phase (see for instance [HöI, Theorem 7.7.1] for the standard version):

**Lemma 5.1.** *Let  $U \subset \mathbb{R}^m$  be open and bounded,  $\tilde{\varphi} \in C^\infty(U; \mathbb{R})$ , and  $a \in C_0^\infty(U)$ . Assume that the following inequalities hold:*

$$\begin{aligned} C_1^{-1} \tilde{h}^{-1} &\leq |d\tilde{\varphi}(x)| \leq C_1 \tilde{h}^{-1} && \text{for all } x \in \text{supp } a; \\ |\partial^\alpha \tilde{\varphi}(x)| &\leq C_{|\alpha|} \tilde{h}^{-\tilde{\rho}|\alpha|} && \text{for all } x \in \text{supp } a \text{ and } |\alpha| \geq 2; \\ |\partial^\alpha a(x)| &\leq C_{|\alpha|} \tilde{h}^{-\tilde{\rho}|\alpha|} && \text{for all } x \in U \text{ and all } \alpha. \end{aligned} \quad (5.11)$$

Here  $\tilde{\rho}, \tilde{h} \in (0, 1)$  and  $C_0, C_1, C_2, \dots$  are positive constants. Then for each  $N \in \mathbb{N}_0$

$$\left| \int_U e^{i\tilde{\varphi}(x)} a(x) dx \right| \leq C'_N \tilde{h}^{N(1-\tilde{\rho})}, \quad (5.12)$$

where the constant  $C'_N$  depends only on  $U, C_0, C_1, \dots, C_{N+1}$ .

**Remark.** Using coordinate charts and a partition of unity for  $a(x)$ , we see that Lemma 5.1 also applies when  $U$  is a manifold; we will typically use it for  $U = \mathbb{S}^{n-1}$ .

*Proof.* Consider the first order differential operator

$$L = -i \sum_{j=1}^m \frac{\partial_j \tilde{\varphi}}{|d\tilde{\varphi}|^2} \partial_j, \quad |d\tilde{\varphi}|^2 := \sum_{j=1}^m |\partial_j \tilde{\varphi}|^2.$$

Then  $e^{i\tilde{\varphi}} = L(e^{i\tilde{\varphi}})$ . Integrating by parts  $N$  times, we obtain

$$\left| \int_U e^{i\tilde{\varphi}(x)} a(x) dx \right| = \left| \int_U e^{i\tilde{\varphi}(x)} (L^t)^N a(x) dx \right| \leq \mu_L(U) \sup_x |(L^t)^N a(x)|,$$

where  $\mu_L$  is the Lebesgue measure and

$$L^t f = i \sum_{j=1}^m \partial_j \left( \frac{\partial_j \tilde{\varphi}}{|d\tilde{\varphi}|^2} f \right).$$

Now, the first two bounds in (5.11) imply that

$$\sup_{x \in \text{supp } a} \left| \partial^\alpha \left( \frac{\partial_j \tilde{\varphi}}{|d\tilde{\varphi}|^2} \right) \right| \leq C'_\alpha \tilde{h}^{1-\tilde{\rho}|\alpha|}$$

where the constants  $C'_\alpha$  depend only on  $C_0, C_1, \dots, C_{|\alpha|+1}$ . This together with the last bound in (5.11) implies the estimate (5.12).  $\square$

Armed with Lemma 5.1, we establish decay of the kernel  $\mathcal{K}$ . We first consider the case when  $y$  and  $y''$  are sufficiently far away from each other so that the corresponding Lagrangian leaves almost do not correlate:

**Lemma 5.2.** *We have uniformly in  $y, y'' \in \mathbb{S}^{n-1}$  such that  $|y - y''| > \frac{1}{2}h^{1/2}$ ,*

$$\mathcal{K}(y, y''; h) = \mathcal{O}(h^\infty).$$

*Proof.* We rewrite (5.10) as

$$\mathcal{K}(y, y''; h) = h^{1-n} \int_{\mathbb{S}^{n-1}} e^{i\varphi(y, y', y'')/h} a(y, y', y''; h) dy', \quad (5.13)$$

$$\varphi = 2(\log |y' - y''| - \log |y' - y|), \quad a = \chi(y', y'') \overline{\chi(y', y)} \psi_0(y'; h).$$

Due to the cutoff  $\chi$ , the amplitude  $a$  is supported inside some fixed compact set which does not intersect  $\{y = y'\}$  and  $\{y' = y''\}$ ; in particular  $\varphi$  is smooth near  $\text{supp } a$ . We next have for all  $N$ ,

$$\|\varphi\|_{C_y^N(\text{supp } a)} \leq C_N |y - y''|,$$

where the constants  $C_N$  are independent of  $y, y'', h$ . Moreover, for some constant  $C$  independent of  $y, y', y'', h$ ,

$$|\partial_{y'} \varphi(y, y', y'')| \geq C^{-1} |y - y''| \quad \text{for } (y, y', y'') \in \text{supp } a.$$

This follows immediately from (1.20) and the fact that the map  $y \mapsto \mathcal{G}(y', y)$  is a diffeomorphism from  $\mathbb{S}^{n-1} \setminus \{y'\}$  to  $T_{y'}^* \mathbb{S}^{n-1}$ .

It remains to apply Lemma 5.1 with

$$\tilde{\varphi} := \frac{\varphi}{h}, \quad \tilde{h} := \frac{h}{|y - y''|} < 2h^{1/2}, \quad \tilde{\rho} := \rho,$$

and use (5.8).  $\square$

Given Lemma 5.2, in order to show (5.9) it suffices to prove the following bound:

$$\sup_{y_0 \in \Lambda_\Gamma} \sup_{y'' \in B(y_0, h^{1/2})} \int_{\Lambda_\Gamma(C_1 h^\rho) \cap B(y_0, h^{1/2})} |\mathcal{K}(y, y''; h)| dy \leq Ch^{2(\beta-\varepsilon)}. \quad (5.14)$$

We claim that it is enough to prove the following  $L^4$  estimate:

$$\sup_{y_0 \in \Lambda_\Gamma} \sup_{y'' \in B(y_0, h^{1/2})} \int_{B(y_0, h^{1/2})} |\mathcal{K}(y, y''; h)|^4 dy \leq Ch^{8(\beta-\varepsilon)-3\rho(n-1-\delta)-\frac{3\delta}{2}}. \quad (5.15)$$

Indeed, (5.14) follows by Hölder's inequality from (5.15) and the following corollary of (5.3):

$$\|\mathbb{1}_{\Lambda_\Gamma(C_1 h^\rho) \cap B(y_0, h^{1/2})}\|_{L^{4/3}} \leq Ch^{\frac{3\rho}{4}(n-1-\delta)+\frac{3\delta}{8}}. \quad (5.16)$$

The proof of (5.15) is based on taking the Taylor expansion of the phase function  $\varphi$  in (5.13) around  $y = y'' = y_0$ . The first term in the expansion is linear in  $y - y''$  and gives the Fourier transform of a distorted version of  $\psi_0$ ; the next terms are  $\mathcal{O}(|y - y_0|^2 + |y'' - y_0|^2) = \mathcal{O}(h)$  and can be put into the amplitude in the integral. The  $L^4$  norm of the Fourier transform can next be estimated via the additive energy of the distorted support of  $\psi_0$ . The proof below relies on this argument, though it does not explicitly use the Fourier transform. Note that to reduce our integral to Fourier transform we needed to restrict to  $y, y'' = y_0 + \mathcal{O}(h^{1/2})$ . To show that the contributions of other  $y, y''$  are negligible in Lemma 5.2 we needed the derivative bounds (5.8), explaining the need to make  $\psi_0$  live on an  $h^{\rho/2}$  neighborhood of the limit set rather than on an  $h^\rho$  neighborhood.

Using Lemma 3.3, take  $\psi_1(y; h) \in C^\infty(\mathbb{S}^{n-1}; [0, 1])$  such that

$$\begin{aligned} \text{supp}(1 - \psi_1) \cap B(y_0, h^{1/2}) &= \emptyset, \quad \text{supp } \psi_1 \subset B(y_0, 2h^{1/2}); \\ \sup_{\mathbb{S}^{n-1}} |\partial^\alpha \psi_1| &\leq C_\alpha h^{-|\alpha|/2}. \end{aligned} \quad (5.17)$$

Then to prove (5.15) is enough to show that

$$\sup_{y_0 \in \Lambda_\Gamma} \sup_{y'' \in B(y_0, h^{1/2})} \int_{\mathbb{S}^{n-1}} \psi_1(y; h) |\mathcal{K}(y, y''; h)|^4 dy \leq Ch^{8(\beta-\varepsilon)-3\rho(n-1-\delta)-\frac{3\delta}{2}}. \quad (5.18)$$

By Fubini's Theorem and (5.10) we have

$$\int_{\mathbb{S}^{n-1}} \psi_1(y; h) |\mathcal{K}(y, y''; h)|^4 dy = \int_{(\mathbb{S}^{n-1})^4} \mathcal{K}_1(y_1, y_2, y_3, y_4, y''; h) dy_1 dy_2 dy_3 dy_4,$$

where

$$\begin{aligned} \mathcal{K}_1 &= h^{4(1-n)} \psi_0(y_1; h) \psi_0(y_2; h) \psi_0(y_3; h) \psi_0(y_4; h) \mathcal{K}_2, \\ \mathcal{K}_2 &= \left( \frac{|y_1 - y''| \cdot |y_3 - y''|}{|y_2 - y''| \cdot |y_4 - y''|} \right)^{2i/h} \chi(y_1, y'') \overline{\chi(y_2, y'')} \chi(y_3, y'') \overline{\chi(y_4, y'')} \mathcal{K}_3, \\ \mathcal{K}_3 &= \int_{\mathbb{S}^{n-1}} \left( \frac{|y_2 - y| \cdot |y_4 - y|}{|y_1 - y| \cdot |y_3 - y|} \right)^{2i/h} \overline{\chi(y_1, y)} \chi(y_2, y) \overline{\chi(y_3, y)} \chi(y_4, y) \psi_1(y; h) dy. \end{aligned}$$

The next statement shows that  $\mathcal{K}_1$  is very small unless  $y_1, y_2, y_3, y_4$  satisfy a certain additive relation. The measure of the set of quadruples  $(y_1, y_2, y_3, y_4)$  which do satisfy this relation will later be estimated using additive energy.

**Lemma 5.3.** *Let  $\eta_j = \mathcal{G}(y_0, y_j) \in T_{y_0}^* \mathbb{S}^{n-1}$ , with  $\mathcal{G}$  defined in (1.19), and assume that*

$$|\eta_1 - \eta_2 + \eta_3 - \eta_4| \geq h^{\rho/2}. \quad (5.19)$$

*Then  $\mathcal{K}_1(y_1, y_2, y_3, y_4, y''; h) = \mathcal{O}(h^\infty)$ , uniformly in  $y_0, y_1, y_2, y_3, y_4, y''$ .*

*Proof.* It is enough to show that  $\mathcal{K}_3 = \mathcal{O}(h^\infty)$ . For that, we write

$$\begin{aligned} \mathcal{K}_3 &= \int_{\mathbb{S}^{n-1}} e^{i\varphi/h} a \, dy, \quad \varphi(y_1, y_2, y_3, y_4, y) = 2 \sum_{j=1}^4 (-1)^j \log |y_j - y|, \\ a(y_1, y_2, y_3, y_4, y; h) &= \overline{\chi(y_1, y)} \chi(y_2, y) \overline{\chi(y_3, y)} \chi(y_4, y) \psi_1(y; h). \end{aligned}$$

Put  $\eta := \eta_1 - \eta_2 + \eta_3 - \eta_4 \in T_{y_0}^* \mathbb{S}^{n-1}$ . By (1.20),

$$\partial_y \varphi(y_1, y_2, y_3, y_4, y_0) = \eta.$$

Since  $\psi_1$  is supported in  $B(y_0, 2h^{1/2})$ , we have for some global constant  $C$ ,

$$|\eta| - Ch^{1/2} \leq |\partial_y \varphi| \leq |\eta| + Ch^{1/2} \quad \text{on } \text{supp } a.$$

By (5.19), we get for  $h$  small enough,

$$|\eta|/2 \leq |\partial_y \varphi| \leq 2|\eta| \quad \text{on } \text{supp } a.$$

It remains to apply Lemma 5.1 with

$$\tilde{\varphi} := \frac{\varphi}{h}, \quad \tilde{h} := \frac{h}{|\eta|} \leq h^{1-\rho/2}, \quad \tilde{\rho} := \frac{1}{2-\rho},$$

and use (5.17). □

Since  $\chi$  is supported away from the diagonal, there exists a constant  $C_1$  such that on  $\text{supp } \mathcal{K}_1$ , we have  $|\mathcal{G}(y_0, y_j)| \leq C_1$  for  $j = 1, 2, 3, 4$ . By (5.7), the additive energy bound (1.21) with  $\alpha = Ch^{\rho/2}$  implies that uniformly in  $y''$ ,

$$\mu_L(\{(y_1, y_2, y_3, y_4) \in \text{supp } \mathcal{K}_1 \mid |\eta_1 - \eta_2 + \eta_3 - \eta_4| \leq h^{\rho/2}\}) \leq Ch^{2\rho(n-1) - \frac{3\rho}{2}\delta + \frac{\rho}{2}\beta_E} \quad (5.20)$$

where  $\eta_j := \mathcal{G}(y_0, y_j) \in T_{y_0}^* \mathbb{S}^{n-1}$ . We also have by (5.17)

$$\sup |\mathcal{K}_1| \leq Ch^{\frac{7}{2}(1-n)}.$$

Together with (5.20) and Lemma 5.3, this gives

$$\sup_{y_0 \in \Lambda_\Gamma} \sup_{y'' \in B(y_0, h^{1/2})} \int_{\mathbb{S}^{n-1}} \psi_1(y; h) |\mathcal{K}(y, y''; h)|^4 \, dy \leq Ch^{(2\rho - \frac{7}{2})(n-1) - \frac{3\rho}{2}\delta + \frac{\rho}{2}\beta_E}.$$

This implies (5.18) as long as

$$\left(2\rho - \frac{7}{2}\right)(n-1) - \frac{3\rho}{2}\delta + \frac{\rho}{2}\beta_E \geq 8(\beta - \varepsilon) - 3\rho(n-1-\delta) - \frac{3\delta}{2}.$$

Recalling (1.22), this inequality becomes

$$\left(5(n-1) - \frac{9}{2}\delta + \frac{1}{2}\beta_E\right)(1-\rho) \leq 8\varepsilon.$$

The last inequality holds when  $\rho$  is close to 1 depending on  $\varepsilon$ , finishing the proof of Theorem 4.

## 6. GENERAL BOUNDS ON ADDITIVE ENERGY

In this section, we prove a new bound (Theorem 6) on the additive energy of general Ahlfors-David regular sets (not just those arising from hyperbolic quotients).

There is substantial conflict of notation between the current section and the rest of the paper. However, this should not cause a problem since the two are completely decoupled from each other. This makes it possible to use simpler notation in the current section.

We first recall the definition of Ahlfors-David regularity:

**Definition 1.4.** *Let  $(\mathcal{M}, d)$  be a complete metric space with more than one element. We say a closed set  $\mathcal{X} \subset \mathcal{M}$  is  $\delta$ -regular with constant  $C_{\mathcal{X}}$  if for all  $x \in \mathcal{X}$  we have*

$$C_{\mathcal{X}}^{-1}r^{\delta} \leq \mu_{\delta}(\mathcal{X} \cap B(x, r)) \leq C_{\mathcal{X}}r^{\delta}, \quad 0 < r < \text{diam}(\mathcal{M})$$

where  $B(x, r)$  is the metric ball of radius  $r$  centered at  $x$  and  $\mu_{\delta}$  is the  $\delta$ -dimensional Hausdorff measure.

**Example 1** (Cantor set). *Let  $\mathcal{X} \subset [0, 1]$  be the middle third Cantor set. Then  $\mathcal{X}$  is  $(\log 2 / \log 3)$ -regular.*

**Example 2** (The limit set of a hyperbolic group). *Let  $\Gamma$  be an  $n$ -dimensional convex co-compact hyperbolic group, let  $\Lambda_{\Gamma} \subset S^{n-1}$  be the limit set, and let  $\delta = \delta(\Gamma)$  be the critical exponent. Then  $\Lambda_{\Gamma}$  is  $\delta$ -regular. See §1.3 and [Su, Theorem 7].*

In this section, it is convenient to use the following definition of additive energy, which is different from Definition 1.2. In §7.2 we will reduce one quantity to the other.

**Definition 6.1** (Additive energy). *Let  $\mathcal{X} \subset [0, 1]^n$  and let  $\mu$  be an outer measure on  $\mathcal{X}$  with  $0 < \mu(\mathcal{X}) < \infty$ . For  $\alpha > 0$ , define the (scale  $\alpha$ ) **additive energy***

$$\mathcal{E}_A(\mathcal{X}, \mu, \alpha) = \mu^4(\{(x_1, x_2, x_3, x_4) \in \mathcal{X}^4 \mid |x_1 - x_2 + x_3 - x_4| < \alpha\}). \quad (6.1)$$

*If the measure  $\mu$  is apparent from context, we will write  $\mathcal{E}_A(\mathcal{X}, \alpha)$  in place of  $\mathcal{E}_A(\mathcal{X}, \mu, \alpha)$ .*

We will now restrict attention to the metric space  $[0, 1]$ . The main result of this section is

**Theorem 6** (Regular sets have small additive energy). *Let  $\mathcal{X} \subset [0, 1]$  be a  $\delta$ -regular set with regularity constant  $C_{\mathcal{X}}$  and  $\delta < 1$ . Then*

$$\mathcal{E}_A(\mathcal{X}, \alpha) = \mathcal{E}_A(\mathcal{X}, \mu_{\delta}, \alpha) \leq \tilde{C} \alpha^{\delta + \beta_{\mathcal{X}}} \quad (6.2)$$

for some  $\beta_{\mathcal{X}} > 0$  and some  $\tilde{C} > 0$ . In particular, we can choose

$$\beta_{\mathcal{X}} = \delta \exp \left[ -\mathbf{K}(1 - \delta)^{-14}(1 + \log^{14} C) \right], \quad (6.3)$$

where  $\mathbf{K}$  is a large absolute constant; the constant  $\tilde{C}$  depends only on  $\delta$  and  $C_{\mathcal{X}}$ .

The exact bound (6.3) is not important and can certainly be improved. The key point is that  $\beta_{\mathcal{X}}$  does not depend on  $\alpha$ .

Heuristically, Theorem 6 says that if we choose points  $x_1, x_2, x_3 \in \mathcal{X}$  at random (using normalized  $\delta$ -dimensional Hausdorff measure on  $\mathcal{X}$ ), then the probability that there will exist a point  $x_4$  with  $|x_1 - x_2 + x_3 - x_4| < \alpha$  is at most  $\tilde{C} \alpha^{\beta_{\mathcal{X}}}$ . For  $\beta_{\mathcal{X}} > 0$  and  $\alpha \ll \tilde{C}^{-1/\beta_{\mathcal{X}}}$ , this quantity is much smaller than 1.

**Remarks.** (i) If  $I \subset \mathbb{R}$  is an interval of finite length and  $\mathcal{X} \subset I$  is a  $\delta$ -regular set, we can apply Theorem 6 to  $\mathcal{X}$  by rescaling  $I$  appropriately. More interestingly, if  $I \subset \mathbb{R}$  is a (possibly infinite) interval and  $\mathcal{X} \subset I$  is a  $\delta$ -regular set, then  $\mathcal{X}' = \mathcal{X} \cap [0, 1]$  might not be  $\delta$ -regular (for example, it might be the union of a  $\delta$ -regular set  $\mathcal{X}_0$  and a point far from  $\mathcal{X}_0$ ), but in many instances we can still apply Theorem 6. This is discussed further in §7.2.

(ii) Theorem 6 is a statement about the additive energy of  $\delta$ -regular sets. In Proposition 6.23 below, we state an alternate version of the theorem that bounds the additive energy of sets that are unions of intervals of length  $\alpha$  (and which satisfy conditions analogous to being  $\delta$ -regular).

The main application of Theorem 6 will be a bound on the additive energy of the limit set of a Fuchsian group. Informally, the result is as follows. Let  $\Lambda_{\Gamma} \subset S^1$  be the limit set of a Fuchsian group with critical exponent  $\delta(\Gamma)$  and let  $\mathcal{G}(y_0, \cdot) : S^1 \rightarrow \mathbb{R}$  be the stereographic projection defined by (1.19). Then the image of  $\Lambda_{\Gamma}$  is a  $\delta(\Gamma)$ -regular set. The set  $\mathcal{G}(y_0, \Lambda_{\Gamma})$  transforms naturally under a certain type of group operation, and this allows us to restrict attention to the interval  $[0, 1]$ . We then apply Theorem 6 to bound the additive energy of  $\mathcal{G}(y_0, \Lambda_{\Gamma}) \cap [0, 1]$ . The exact statement and proof are in §7.

### 6.1. Ideas behind the proof of Theorem 6.

6.1.1. *Ahlfors-David regularity and arithmetic progressions.* A  $\delta$ -regular subset of  $[0, 1]$  cannot contain long arithmetic progressions. More precisely, suppose  $P \subset \mathcal{X}$  is an arithmetic progression of length  $|P|$  and spacing  $t > 0$ . Let  $I \subset [0, 1]$  be the interval of length  $t|P|$  centered around  $P$ . If we place an interval of radius  $t/2$  around each point of  $P$ , then  $\mu_\delta(\mathcal{X} \cap I) \geq C_\mathcal{X}^{-1}|P|(t/2)^\delta$ . On the other hand,  $\mu_\delta(\mathcal{X} \cap I) \leq C_\mathcal{X}(|P|t)^\delta$ . If  $|P|$  is sufficiently large (depending on  $C_\mathcal{X}$  and  $\delta$ ), we arrive at a contradiction, provided that  $\delta < 1$ . In fact more is true. If  $P$  is not contained in  $\mathcal{X}$  but merely meets  $\mathcal{X}$  in many points, the argument still applies as well. Finally, the argument is not affected if we perturb the points of  $P$  slightly. We say that  $\mathcal{X}$  *strongly avoids long arithmetic progressions*.

Note that this argument relies on the fact that  $\delta < 1$ . If instead  $\mathcal{X}$  is a subset of  $[0, 1]^n$  for  $n > 1$  (or a more general metric space) and  $\delta \geq 1$  then the argument fails. In §6.8.2 we will discuss this phenomenon further.

6.1.2. *Small doubling and additive structure.* If  $A \subset \mathbb{Z}$  is a finite set and  $|A + A| < K|A|$ , what can we say about  $A$ ? There is a family of theorems in additive combinatorics that say that if  $K$  is small then  $A$  must have additive structure. The most famous of these is Freïman's theorem [Fr], which says that  $A$  must be contained in a generalized arithmetic progression. For our purposes however, we will obtain stronger results by using a variant of Freïman's theorem due to Sanders [Sa], which makes the weaker claim that  $A$  has large intersection with a convex progression.

When combined with the ideas discussed above, we conclude that if  $\mathcal{X}$  is a regular set then  $\mathcal{X}$  cannot have maximally large additive energy. Unfortunately, the sort of bounds that one obtains from this argument are very weak—far too weak to obtain the polynomial in  $\alpha$  improvement of Theorem 6<sup>†</sup>.

6.1.3. *Multiscale analysis of Ahlfors-David regular sets.* If  $\mathcal{X}$  is a regular set, we can examine  $\mathcal{X}$  at many intermediate scales between  $\alpha$  and 1; there will be roughly  $|\log \alpha|$  scales total. We use the arguments above to get a small gain in the scale- $\alpha$  additive energy of  $\mathcal{X}$  at each intermediate scale. These gains will compound with each intermediate scale, and the total gain will be large enough to obtain Theorem 6.

6.2. **Ahlfors-David regular sets and additive structure.** We start the proof of Theorem 6 by exploring the implications of  $\delta$ -regularity for the additive structure of the set  $\mathcal{X}$ .

**Definition 6.2.** *An arithmetic progression is a set of the form*

$$\{a - \ell q, a - (\ell - 1)q, \dots, a, a + q, \dots, a + \ell q\} \subset \mathbb{Z},$$

---

<sup>†</sup>However, if the polynomial Freïman-Ruzsa conjecture is proved then this theorem may be employed directly, and the subsequent steps would not be needed.

where  $a, q \in \mathbb{Z}$ ,  $q \neq 0$ , and  $\ell \geq 0$ .

**Definition 6.3.** Let  $A \subset \mathbb{Z}$  be a finite set. We say that  $A$  **strongly avoids long arithmetic progressions** (with parameter  $S$ ) if for all  $\varepsilon > 0$ , there is a number  $S = S(\varepsilon)$  so that for all arithmetic progressions  $P \subset \mathbb{Z}$  with  $|P \cap A| \geq \varepsilon|P|$ , we have  $|P| \leq S(\varepsilon)$ .

If  $\mathcal{X} \subset t\mathbb{Z}$  for  $0 < t < 1$ , then we say that  $\mathcal{X}$  strongly avoids long arithmetic progressions if  $t^{-1}\mathcal{X}$  does, i.e. we simply re-scale  $\mathcal{X}$  so that it lies on the integer lattice.

**Definition 6.4.** A ( $d$ -dimensional) **centered convex progression** is a triple  $(B, \Lambda, \varphi)$ , where  $B \subset \mathbb{R}^d$  is a centrally symmetric convex set,  $\Lambda \subset \mathbb{R}^d$  is a lattice, and  $\varphi : \Lambda \rightarrow \mathbb{Z}$  is a linear map. We will primarily be interested in the image  $\varphi(B \cap \Lambda)$ . Sets of this form are generalizations of arithmetic progressions. If  $\varphi$  is injective on  $B \cap \Lambda$ , we say that  $(B, \Lambda, \varphi)$  is **proper**.

The next lemma shows that a  $d$ -dimensional centered convex progression  $(B, \Lambda, \varphi)$  of cardinality  $N := |\varphi(B \cap \Lambda)|$  can be embedded into a centered convex progression  $(B', \Lambda', \varphi')$  whose size  $|B' \cap \Lambda'|$  is bounded by  $N$  times a  $d$ -dependent constant.

**Lemma 6.5** (Cardinality vs. size). Let  $(B, \Lambda, \varphi)$  be a  $d$ -dimensional centered convex progression. Then there is some  $d' \leq d$  and a  $d'$ -dimensional centered convex progression  $(B', \Lambda', \varphi')$  with

$$|B' \cap \Lambda'| \leq 2^{(d+2)^2} |\varphi(B \cap \Lambda)| \quad (6.4)$$

such that

$$\varphi(B \cap \Lambda) \subset \varphi'(B' \cap \Lambda'). \quad (6.5)$$

*Proof.* We recall [TaVu08, Corollary 4.2]:

**Proposition 6.6.** Let  $(B, \Lambda, \varphi)$  be a  $d$ -dimensional centered convex progression. Then there exists a  $d'$ -dimensional **proper** centered convex progression  $(B'', \Lambda'', \varphi'')$  for some  $d' \leq d$  such that we have the inclusions

$$\varphi(B \cap \Lambda) \subset \varphi''((2^{d-d'+1}B'') \cap \Lambda''), \quad (6.6)$$

$$\varphi((2B'') \cap \Lambda'') \subset \varphi(B \cap \Lambda), \quad (6.7)$$

where  $tB = \{x \in \mathbb{R}^d \mid t^{-1}x \in B\}$ .

Let  $B' := 2^{d-d'+1}B''$ ,  $\Lambda' := \Lambda''$ , and  $\varphi' := \varphi''$ . Then (6.5) is satisfied. Since  $(B'', \Lambda'', \varphi'')$  is proper, we have

$$|(2^{-d+d'-1}B') \cap \Lambda'| = |\varphi'((2^{-d+d'-1}B') \cap \Lambda')| \leq |\varphi(B \cap \Lambda)|. \quad (6.8)$$

By [TaVu08, Lemma 3.3], we have

$$|B' \cap \Lambda'| \leq 2^{(d+2)^2} |(2^{-d+d'-1}B') \cap \Lambda'|. \quad (6.9)$$

Combining (6.8) and (6.9) we obtain (6.4).  $\square$

We next recall [TaVu06, Lemma 3.36]:

**Proposition 6.7.** *Let  $B \subset \mathbb{R}^d$  be a centrally symmetric convex set and let  $\Lambda \subset \mathbb{R}^d$  be a lattice. Suppose that the  $\mathbb{R}$  span of  $B \cap \Lambda$  has dimension  $r$ . Then there exists an  $r$ -tuple  $w = (w_1, \dots, w_r)$  with  $w_1, \dots, w_r$  linearly independent vectors in  $\Lambda$ , and an  $r$ -tuple of integers  $N = (N_1, \dots, N_r)$  so that*

$$[-N, N] \cdot w \subset B \cap \Lambda \subset [-r^{2r}N, r^{2r}N] \cdot w. \quad (6.10)$$

Here

$$[-N, N] \cdot w = \{\ell_1 w_1 + \dots + \ell_r w_r \mid -N_j \leq \ell_j \leq N_j, j = 1, \dots, r\}.$$

**Corollary 6.8.** *Let  $B \subset \mathbb{R}^d$  be a centrally symmetric convex set and let  $\Lambda \subset \mathbb{R}^d$  be a lattice. Suppose that the  $\mathbb{R}$  span of  $B \cap \Lambda$  has dimension  $r$ . Then there exists an  $r$ -tuple  $w = (w_1, \dots, w_r)$  with  $w_1, \dots, w_r$  linearly independent vectors in  $\Lambda$ , and an  $r$ -tuple of integers  $N = (N_1, \dots, N_r)$  so that*

$$B \cap \Lambda \subset [-N, N] \cdot w, \quad (6.11)$$

and

$$\prod_{j=1}^r (2N_j + 1) \leq 3^r r^{2r^2} |B \cap \Lambda|. \quad (6.12)$$

**Lemma 6.9** (Arithmetic progressions inside convex progressions). *Let  $(B, \Lambda, \varphi)$  be a  $d$ -dimensional centered convex progression and let*

$$A \subset \varphi(B \cap \Lambda), \quad |A| \geq \varepsilon |\varphi(B \cap \Lambda)|.$$

Then there exists a (one dimensional) arithmetic progression  $P \subset \mathbb{Z}$  with

$$|P| \geq |A|^{1/d} \quad (6.13)$$

and

$$|P \cap A| \geq d^{-3(d+2)^2} \varepsilon |P|. \quad (6.14)$$

*Proof.* Let  $(B', \Lambda', \varphi')$  be a centered convex progression obeying (6.4) and (6.5). In particular, we have  $A \subset \varphi'(B' \cap \Lambda')$  and

$$|B' \cap \Lambda'| \leq 2^{(d+2)^2} \varepsilon^{-1} |A|. \quad (6.15)$$

By Corollary 6.8, there is a number  $r \leq d$ , an  $r$ -tuple  $w = (w_1, \dots, w_r)$  of linearly independent vectors in  $\Lambda'$ , and an  $r$ -tuple of integers  $N = (N_1, \dots, N_r)$  so that

$$(\varphi')^{-1}(A) \cap B' \cap \Lambda' \subset B' \cap \Lambda' \subset [-N, N] \cdot w, \quad (6.16)$$

and

$$\prod_{j=1}^r (2N_j + 1) \leq 3^r r^{2r^2} |B' \cap \Lambda'| \leq 2^{(d+3)^2} d^{2d^2} \varepsilon^{-1} |A|. \quad (6.17)$$

Now, (6.16) implies that

$$A \subset \varphi'([-N, N] \cdot w). \quad (6.18)$$

We can assume that  $\varphi'(w_i) \neq 0$  for each index  $i$  for which  $N_i \neq 0$ , since if  $\varphi'(w_i) = 0$  then we could set  $N_i = 0$  and both (6.17) and (6.18) remain satisfied.

By re-indexing if necessary, we can assume that  $N_1 \geq N_j$  for all  $j = 2, \dots, r$ . Partition the set  $[-N, N] \cdot w$  into  $\prod_{j=2}^r (2N_j + 1)$  disjoint sets of the form

$$\left\{ \ell w_1 + \sum_{j=2}^r x_j w_j \mid -N_1 \leq \ell \leq N_1 \right\}.$$

Each of these sets has cardinality

$$(2N_1 + 1) \geq \left( \prod_{j=1}^r (2N_j + 1) \right)^{1/r} \geq |A|^{1/d}.$$

By (6.17), (6.18), and pigeonholing, at least one of these sets  $Y$  must satisfy

$$|Y \cap (\varphi')^{-1}(A)| \geq 2^{-(d+3)^2} d^{-2d^2} \varepsilon |Y| \geq d^{-3(d+2)^2} \varepsilon |Y|. \quad (6.19)$$

For the last inequality we may assume that  $d \geq 2$ , since otherwise we may put  $P := \varphi(B \cap \Lambda)$ .

Let  $P = \varphi'(Y)$ . By assumption  $\varphi'$  is injective on  $Y$ , so  $P$  satisfies (6.13) and (6.14).  $\square$

The following Proposition is a direct corollary of [Sa, Theorem 1.4]:

**Proposition 6.10** (Additive structure). *Let  $A \subset \mathbb{Z}$  and suppose  $|A + A| \leq K|A|$ . Then there is a  $d \leq K_0 \log^4 K$  dimensional centered convex progression  $(B, \mathbb{Z}^d, \varphi)$  and an offset  $x \in \mathbb{Z}$  so that*

$$|\varphi(B \cap \mathbb{Z}^d)| \leq \exp[K_0 \log^4 K] \cdot |A| \quad (6.20)$$

and

$$|(A - x) \cap \varphi(B \cap \mathbb{Z}^d)| \geq \exp[-K_0 \log^4 K] \cdot |A|. \quad (6.21)$$

Here  $K_0 > 0$  is an absolute constant.

Applying Lemma 6.9, we obtain the following corollary.

**Corollary 6.11.** *Let  $A \subset \mathbb{Z}$  and suppose that  $|A + A| \leq K|A|$ . Then there is an arithmetic progression  $P$  so that*

$$|P| \geq |A|^{(K_1 \log^4 K)^{-1}}, \quad (6.22)$$

and

$$|A \cap P| \geq e^{-K_1 \log^9 K} |P|. \quad (6.23)$$

Here  $K_1 > 0$  is an absolute constant. The term  $\log^9$  in (6.23) could be replaced with  $\log^{8+o(1)}$ , but we will not worry about these small optimizations.

**Corollary 6.12.** *Let  $A \subset \mathbb{Z}$ . Suppose that  $A$  strongly avoids long arithmetic progressions (with parameter  $S(\varepsilon)$ ), and  $|A + A| \leq K|A|$ . Then*

$$|A| \leq \left( S(e^{-K_1 \log^9 K}) \right)^{K_1 \log^4 K}, \quad (6.24)$$

where  $K_1$  is the absolute constant from Corollary 6.11.

**Proposition 6.13** (Ahlfors-David regular sets avoid arithmetic progressions). *Let  $\mathcal{X} \subset [0, 1]$  be a  $\delta$ -regular set with regularity constant  $C$  and let  $0 < \alpha < 1$ . Then  $\mathcal{X}(\alpha) \cap \alpha\mathbb{Z}$  strongly avoids long arithmetic progressions (here  $\mathcal{X}(\alpha)$  is the  $\alpha$ -neighborhood of  $\mathcal{X}$ ). In particular, we can take  $S(\varepsilon) = (10C^2\varepsilon^{-1})^{\frac{1}{1-\delta}}$ .*

*Proof.* Let  $P \subset [0, 1] \cap \alpha\mathbb{Z}$  be a proper arithmetic progression. In particular,  $P$  has spacing  $t \geq \alpha$ . Assume that  $|P \cap \mathcal{X}(\alpha)| > \varepsilon|P|$ . For each point  $y \in P \cap \mathcal{X}(\alpha)$ , the ball  $B(y, 2t)$  contains the ball of radius  $t$  centered at some point of  $\mathcal{X}$ . By (1.23) we have

$$\sum_{y \in P \cap \mathcal{X}(\alpha)} \mu_\delta(\mathcal{X} \cap B(y, 2t)) \geq C^{-1}\varepsilon|P|t^\delta. \quad (6.25)$$

On the other hand, the balls  $B(y, 2t)$  are at most five-fold overlapping, and  $P$  is contained in an interval  $J$  of length  $(|P| + 3)t \leq 2|P|t$  (unless  $|P| < 3$  in which case  $|P| < S(\varepsilon)$  automatically). By (1.23) we have

$$\sum_{y \in P \cap \mathcal{X}(\alpha)} \mu_\delta(\mathcal{X} \cap B(y, 2t)) \leq 5\mu_\delta(\mathcal{X} \cap J) \leq 10C(t|P|)^\delta. \quad (6.26)$$

We conclude that  $|P| \leq (10C^2\varepsilon^{-1})^{\frac{1}{1-\delta}}$ . □

Combining Corollary 6.12 and Proposition 6.13, we get

**Corollary 6.14.** *Let  $\mathcal{X} \subset [0, 1]$  be a  $\delta$ -regular set with regularity constant  $C$ . Let  $0 < \alpha < 1$  and  $A \subset \mathcal{X}(\alpha) \cap \alpha\mathbb{Z}$ , and suppose  $|A + A| \leq K|A|$ . Then*

$$|A| \leq \exp \left[ K_3(1 - \delta)^{-1}(\log C) \log^{13} K \right] \quad (6.27)$$

for some absolute constant  $K_3$ .

**6.3. Ahlfors-David regular trees.** The key to proving Theorem 6 will be to analyze  $\mathcal{X}$  at many scales. Heuristically, if  $\alpha = M^{-N}$  for  $M, N$  positive integers, then  $\mathcal{X}$  can naturally be analyzed at the  $N$  scales  $1, M^{-1}, \dots, M^{-N}$ . On each scale, we will get a small gain in the scale- $\alpha$  additive energy of  $\mathcal{X}$ . In order to keep track of  $\mathcal{X}(\alpha)$  at these different scales we will construct an object called a tree.

**Definition 6.15** (Trees). *A (**rooted**) tree of height  $\ell \in \mathbb{Z}_{\geq 0}$  is a connected acyclic graph with a distinguished vertex (called the root). Once we have specified the root of a tree, each vertex has a well-defined height (i.e. its distance from the root), and we say*

that one vertex  $v$  is a parent of another vertex  $v'$  if  $v$  and  $v'$  are adjacent and  $v$  has smaller height.

More formally, a (rooted) tree is a quadruple  $(V, H, p, \ell)$ , where

- $V$  is a finite set of **vertices**;
- $H : V \rightarrow \{0, \dots, \ell\}$  is the **height** function, and  $H^{-1}(0)$  consists of a single vertex called the **root**;
- $p : V \setminus H^{-1}(0) \rightarrow V$  is the **parent** function, and  $p(H^{-1}(t)) \subset H^{-1}(t-1)$  for  $t = 1, \dots, \ell$ .

We denote by  $\mathcal{V}_t(T) := H^{-1}(t)$  the set of vertices of height  $t$ .

Let  $T$  be a tree of height  $\ell$ . The set of *leaves* of  $T$  is defined as

$$\mathcal{L}(T) = H^{-1}(\ell) \subset V.$$

For each vertex  $v \in V$ , we say that  $v' \in V$  is a *child* of  $v$ , if  $p(v') = v$ . More generally, we say that  $v'$  is *below*  $v$ , and write  $v' \prec v$ , if there is a sequence of vertices  $v_1, \dots, v_m$  so that  $v_1 = v$ ,  $v_m = v'$ , and  $v_{i+1}$  is a child of  $v_i$  for each  $i = 1, \dots, m-1$ .

If  $T$  is a tree and  $v \in V$ , we define  $T_v$  to be the *subtree* of  $T$  rooted at  $v$ . This is a tree of height  $\ell - H(v)$ . Its vertices are the set  $\{v' \in V \mid v' \prec v\}$ . The height function is  $v' \mapsto H(v') - H(v)$ , and the parent function is inherited from the original tree.

**Definition 6.16** (Regular trees). *Let  $\ell, B, C \geq 0$ . We say a tree  $T$  is an “Ahlfors-David regular tree of height  $\ell$ , branching  $B$ , and regularity constant  $C$ ” if  $T$  is a tree of height  $\ell$  and for each  $v \in T$ ,*

$$C^{-1}B^{\ell-H(v)} \leq |\mathcal{L}(T_v)| \leq CB^{\ell-H(v)}, \quad (6.28)$$

where  $T_v$  is the sub-tree of  $T$  rooted at  $v$ .

**Remark.** If  $T$  is an Ahlfors-David regular tree with height  $\ell$ , branching  $B$ , and regularity constant  $C$ , then each vertex of  $T$  has between  $C^{-2}B$  and  $C^2B$  children. However, much more is true—if a vertex of  $T$  has (relatively) few children, then *its* children must have many children, and vice versa. Thus the tree  $T$  might not be perfectly balanced, but it cannot become extremely unbalanced either.

**Lemma 6.17.** *Let  $T$  be an Ahlfors-David regular tree with height  $\ell$ , branching  $B$ , and regularity constant  $C$ . Let  $v \in T$ . Then  $T_v$  (the sub-tree rooted at  $v$ ) is an Ahlfors-David regular tree of height  $\ell - H(v)$ , branching  $B$ , and regularity constant  $C$ .*

If  $T$  is a tree of height  $\ell$  and  $j \in \mathbb{N}$ , then we can define the  $j$ -th *power* of  $T$ , denoted  $T^j$ , as the following tree of height  $\ell$ :

- the vertices of  $T^j$  are ordered pairs  $(v_1, \dots, v_j)$ , where  $v_1, \dots, v_j$  are vertices of the tree  $T$  of the same height;

- the height of a vertex  $(v_1, \dots, v_j)$  is equal to the height of each of  $v_i$ ;
- the parent of a vertex  $(v_1, \dots, v_j)$  is equal to  $(p(v_1), \dots, p(v_j))$ , where  $p$  is the parent function of the original tree.

If  $T$  is an Ahlfors-David regular tree with height  $\ell$ , branching  $B$ , and regularity constant  $C$ , then  $T^j$  is an Ahlfors-David regular tree with height  $\ell$ , branching  $B^j$ , and regularity constant  $C^j$ .

**6.4. Discretization.** The trees discussed in the previous section are useful for describing the multi-scale structure of  $\delta$ -regular sets.

Let  $\mathcal{X} \subset [0, 1]$  be a  $\delta$ -regular set with regularity constant  $C$ . Define

$$C_1 := (10C^2)^{\frac{1}{1-\delta}}. \quad (6.29)$$

Let  $M, N$  be positive integers; we will fix  $M$  and study asymptotic behavior as  $N \rightarrow \infty$ . We will describe a process that divides  $[0, 1]$  into sub-intervals of length roughly  $M^{-j}$ ,  $j = 0, \dots, N$ , and assembles these intervals into a tree.

For each  $j = 0, \dots, N$ , divide  $[0, 1]$  into  $M^j$  intervals of the form  $[iM^{-j}, (i+1)M^{-j})$ . If  $I$  is an interval of this form, we say  $I$  is *empty* if  $I \cap \mathcal{X} = \emptyset$ . Otherwise  $I$  is *non-empty*. If several non-empty intervals are adjacent, merge them into a single (longer) interval. By Lemma 6.13 with  $\alpha := M^{-j}$  and  $\varepsilon := 1$ , at most  $C_1$  intervals can be merged into a single interval in this fashion. Let  $\mathcal{I}_j$  be the set of non-empty intervals obtained after the merger process is complete. Each interval in  $\mathcal{I}_j$  has length between  $M^{-j}$  and  $C_1 M^{-j}$ . Furthermore, if  $I$  is an interval in  $\mathcal{I}_j$ , then there is a gap of size  $\geq M^{-j}$  on either side of  $I$  that is disjoint from  $\mathcal{X}$ .

**Lemma 6.18.** *Let  $I \in \mathcal{I}_j$ . Then there is a unique interval  $\tilde{I} \in \mathcal{I}_{j-1}$  that intersects  $I$ . Furthermore,  $I \subset \tilde{I}$ .*

*Proof.* First, note that  $I \subset \bigcup_{\tilde{I} \in \mathcal{I}_{j-1}} \tilde{I}$ . Thus it suffices to show that  $I$  intersects at most one interval from  $\mathcal{I}_{j-1}$ . Suppose that  $I$  intersects two intervals,  $\tilde{I} = [\tilde{i}_0 M^{1-j}, \tilde{i}_1 M^{1-j})$  and  $\tilde{I}' = [\tilde{i}'_0 M^{1-j}, \tilde{i}'_1 M^{1-j})$  from  $\mathcal{I}_{j-1}$ . If we write  $I = [i_0 M^{-j}, i_1 M^{-j})$ , then  $i_0 \leq M\tilde{i}_1 \leq M\tilde{i}'_0 \leq i_1$ .

Since no two intervals in  $\mathcal{I}_{j-1}$  can be adjacent, there must be some interval of the form  $[i' M^{1-j}, (i'+1)M^{1-j})$  that is disjoint from  $\mathcal{X}$ , with  $\tilde{i}_1 < i' < i'+1 < \tilde{i}'_0$ . This implies that  $[i' M^{1-j}, (i'+1)M^{1-j}) \subset I$ , and in particular,  $[(Mi')M^{-j}, (Mi'+1)M^{-j}) \subset I$ . But this implies that  $[(Mi')M^{-j}, (Mi'+1)M^{-j}) \cap I \cap \mathcal{X} = \emptyset$ , which is a contradiction—by the construction of the intervals in  $\mathcal{I}_j$ , every sub-interval of  $I$  of the form  $[iM^{-j}, (i+1)M^{-j})$  must intersect  $\mathcal{X}$ . We conclude that there is at most one interval from  $\mathcal{I}_{j-1}$  that intersects  $I$ .  $\square$

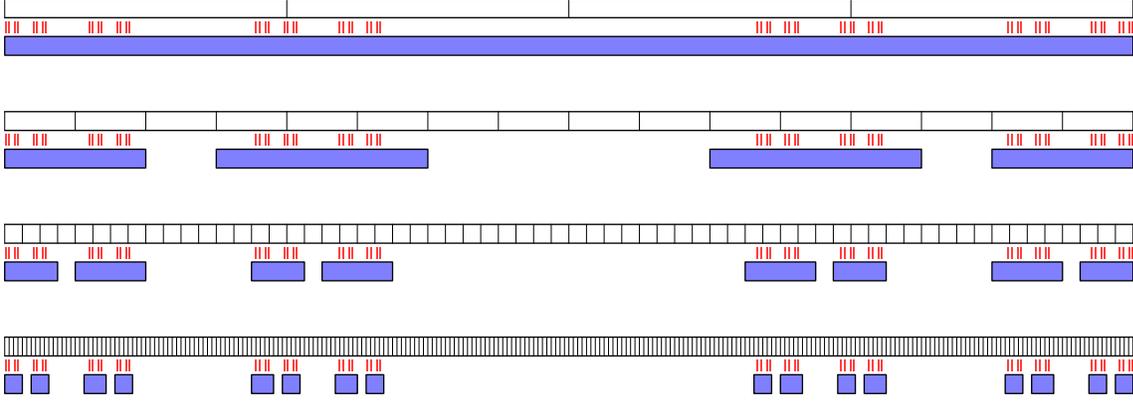


FIGURE 7. The discretization of the middle third Cantor set (pictured in red) with  $M = 4$ . The intervals of the discretization for  $j = 1, \dots, 4$  are pictured in blue and form a tree.

We now construct the tree  $T_{\mathcal{X};M,N}$  as follows. The root vertex of  $T_{\mathcal{X};M,N}$  corresponds to the interval  $[0, 1) \in \mathcal{I}_0$ . For each  $j = 1, \dots, N$ , the vertices of  $T_{\mathcal{X};M,N}$  of height  $j$  correspond to the intervals in  $\mathcal{I}_j$ . The parent of an interval in  $\mathcal{I}_j$  is the unique interval in  $\mathcal{I}_{j-1}$  containing that interval. If  $v$  is a vertex of  $T_{\mathcal{X};M,N}$ , let  $I_v$  be the corresponding interval. Note that  $v' \prec v$  if and only if  $I_{v'} \subset I_v$ . See Figure 7.

**Lemma 6.19.** *Let  $\mathcal{X}$  be a  $\delta$ -regular set with regularity constant  $C$ . Then  $T_{\mathcal{X};M,N}$  is an Ahlfors-David regular tree with height  $N$ , branching  $M^\delta$  and regularity constant*

$$C_2 := C^2 C_1^\delta = 10^{\frac{\delta}{1-\delta}} C^{\frac{2}{1-\delta}}. \quad (6.30)$$

*Proof.* Let  $v$  be a vertex of  $T = T_{\mathcal{X};M,N}$  and let  $I_v \subset [0, 1)$  be the corresponding interval. Choose any  $x_0 \in I_v \cap \mathcal{X}$ . Since  $I_v$  has length no more than  $C_1 M^{-H(v)}$ , it is contained in the ball  $B(x_0, C_1 M^{-H(v)})$ . On the other hand,  $^\ddagger B(x_0, M^{-H(v)}) \cap \mathcal{X} \subset I_v$ . Since  $\mathcal{X}$  is  $\delta$ -regular, we have

$$C^{-1} M^{-H(v)\delta} \leq \mu_\delta(\mathcal{X} \cap I_v) \leq C(C_1 M^{-H(v)})^\delta. \quad (6.31)$$

For each leaf  $v' \in \mathcal{L}(T)$  with  $v' \prec v$ , let  $I_{v'}$  be the corresponding interval. Again, we have

$$C^{-1} M^{-N\delta} \leq \mu_\delta(\mathcal{X} \cap I_{v'}) \leq C(C_1 M^{-N})^\delta. \quad (6.32)$$

On the other hand, by Lemma 6.18 we have

$$\mu_\delta(\mathcal{X} \cap I_v) = \sum_{v' \in \mathcal{L}(T_v)} \mu_\delta(\mathcal{X} \cap I_{v'}).$$

<sup>‡</sup>This is one of the places where we use the merging of consecutive intervals; otherwise, one of the intervals may intersect  $\mathcal{X}$  at an endpoint and the resulting tree may not be regular.

It follows that

$$C_2^{-1}(M^\delta)^{N-H(v)} \leq |\mathcal{L}(T_v)| \leq C_2(M^\delta)^{N-H(v)}, \quad C_2 = C^2 C_1^\delta. \quad \square$$

We will mainly be interested in the tree  $T_{\mathcal{X};M,N}^3$ . A vertex of this tree is a triple of intervals  $(I_1, I_2, I_3)$  where each interval  $I_i$  meets  $\mathcal{X}$  and all three intervals lie in  $\mathcal{I}_j$  for some  $j$  (and thus they have comparable lengths).

**6.5. Additive structure and pruning the tree.** Let  $\mathcal{X}$  be a  $\delta$ -regular set with regularity constant  $C$  and let  $T_{\mathcal{X};M,N}$  be the Ahlfors-David regular tree described in §6.4. Let  $v = (I_1, I_2, I_3)$  be a vertex of  $T_{\mathcal{X};M,N}^3$ . Consider the interval

$$I_1 - I_2 + I_3 := \{x_1 - x_2 + x_3 \mid x_1 \in I_1, x_2 \in I_2, x_3 \in I_3\}.$$

We say that  $v$  *misses*  $\mathcal{X}$  if  $(I_1 - I_2 + I_3) \cap \mathcal{X}(M^{-H(v)}) = \emptyset$ . Otherwise  $v$  *hits*  $\mathcal{X}$ .

The following result shows that if  $M$  is sufficiently large then we have an improvement in additive energy on each level of the tree.

**Proposition 6.20.** *Let  $\mathcal{X}$  be a  $\delta$ -regular set with regularity constant  $C$ , and let  $T_{\mathcal{X};M,N}$  be the Ahlfors-David regular tree described above. Assume that  $M \geq M_0$ , where*

$$M_0 := \exp [K_5 \delta^{-1} (1 - \delta)^{-14} (1 + \log^{14} C)], \quad (6.33)$$

and  $K_5$  is a large absolute constant. Let  $v \in T_{\mathcal{X};M,N}^3$  be a vertex that is not a leaf. Then at least one of the children of  $v$  misses  $\mathcal{X}$ .

*Proof.* Suppose  $M \geq M_0$ . Let  $v \in T_{\mathcal{X};M,N}^3$  be a vertex that is not a leaf and suppose all of the children of  $v$  hit  $\mathcal{X}$ ; we will obtain a contradiction. Write  $v = (v_1, v_2, v_3)$  and let  $I_1, I_2, I_3 \in \mathcal{I}_{H(v)}$  be the corresponding triple of intervals. Let  $\alpha_1 = M^{-H(v)-1}$  and define

$$\begin{aligned} A_i &:= I_i \cap \mathcal{X}(\alpha_1) \cap \alpha_1 \mathbb{Z}, \quad i = 1, 2, 3; \\ A_4 &:= (I_1 - I_2 + I_3) \cap \mathcal{X}(4C_1 \alpha_1) \cap \alpha_1 \mathbb{Z}. \end{aligned}$$

Here  $C_1$  is defined in (6.29).

**Lemma 6.21.** *If every child of  $v$  hits  $\mathcal{X}$ , then*

$$A_1 - A_2 + A_3 \subset A_4. \quad (6.34)$$

*Proof.* Take  $a_i \in A_i$ ,  $i = 1, 2, 3$ . Then  $|a_i - b_i| \leq \alpha_1$  for some  $b_i \in \mathcal{X}$ ; since  $I_i$  is surrounded by size  $M\alpha_1$  intervals which do not intersect  $\mathcal{X}$ , we have  $b_i \in I_i \cap \mathcal{X}$ . Then  $b_i$  lies in some child  $I'_i$  of  $I_i$ , which is an interval of size no more than  $C_1 \alpha_1$ . We have  $I'_1 - I'_2 + I'_3 \cap \mathcal{X}(\alpha_1) \neq \emptyset$  and thus  $b_1 - b_2 + b_3 \in \mathcal{X}((3C_1 + 1)\alpha_1)$ . Then  $a_1 - a_2 + a_3 \in \mathcal{X}(4C_1 \alpha_1)$  and (6.34) follows.  $\square$

We next claim that

$$\begin{aligned} |A_i| &\geq (2C^2)^{-1}M^\delta, \quad i = 1, 2, 3; \\ |A_4| &\leq 44C^2C_1M^\delta. \end{aligned} \tag{6.35}$$

These inequalities follow immediately from (1.23) and the following observations:

- The balls  $B(a, \alpha_1)$  centered at  $a \in A_i$  cover  $\mathcal{X} \cap I_i$ , and  $\mathcal{X} \cap I_i$  contains the intersection of  $\mathcal{X}$  with a ball of radius  $M\alpha_1$  centered at a point of  $\mathcal{X}$ .
- For each  $a \in A_4$ , the ball  $B(a, 5C_1\alpha_1)$  contains a ball of radius  $C_1\alpha_1$  centered at a point of  $\mathcal{X}$ , the balls  $B(a, 5C_1\alpha_1)$  for different  $a$  have overlapping at most  $11C_1$ , and their union is contained in an interval of length  $(3M + 10)C_1\alpha_1 \leq 4MC_1\alpha_1$ .

We now use the Ruzsa sum inequality [Ru, (4.6)] (see also Petridis [Pe]):

$$|A + C| \leq \frac{|A + B| \cdot |B + C|}{|B|}$$

valid for nonempty finite sets  $A, B, C \subset \mathbb{Z}$ . Putting  $A := A_1, B := A_3, C := A_1$  and using (6.35) and the fact that  $|A_1 + A_3| \leq |A_4|$  we obtain

$$|A_1 + A_1| \leq \frac{|A_4|^2}{|A_1| \cdot |A_3|} \cdot |A_1| \leq C_3|A_1|,$$

where

$$C_3 := 7744C^8C_1^2 \leq (10C^2)^{\frac{6}{1-\delta}}.$$

Apply Corollary 6.14 to the set  $A_1$  with  $K := C_3$ . We conclude that

$$|A_1| \leq \exp [K_4(1 - \delta)^{-1}(\log C)((1 - \delta)^{13}(1 + \log C)^{13})], \tag{6.36}$$

where  $K_4$  is an absolute constant. If  $M \geq M_0$ , we obtain a contradiction with the first bound in (6.35).  $\square$

**6.6. Analyzing the pruned tree.** To take advantage of Proposition 6.20, we prove the following general fact about pruned subtrees of Ahlfors–David regular trees.

For two trees  $T = (V, H, p, \ell), T' = (V', H', p', \ell)$  of same height, we say that  $T'$  is a *subtree* of  $T$  if  $V' \subset V$  and  $H', p'$  are the restrictions of  $H, p$  to  $V'$ . We say that  $T'$  is a *pruned subtree*, if for each  $v \in V'$  which is not a leaf, there exists a child of  $v$  in  $T$  which does not lie in  $V'$ . See Figure 8.

**Lemma 6.22** (Pruned trees have few leaves). *Let  $T$  be an Ahlfors–David regular tree with height  $\ell$ , branching  $B$ , and regularity constant  $C$ . Let  $T'$  be a pruned subtree of  $T$ . Then*

$$|\mathcal{L}(T')| \leq (1 - C^{-2}B^{-1})^\ell |\mathcal{L}(T)|. \tag{6.37}$$

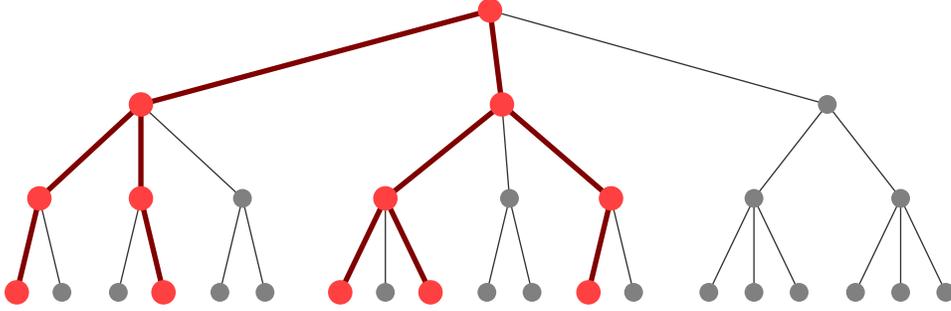


FIGURE 8. An Ahlfors-David regular tree and a pruned subtree (in red).

*Proof.* We will prove the lemma by induction on  $\ell$ , the height of the tree. If  $\ell = 0$  the result is trivial. Now assume the result has been proved for all Ahlfors-David regular trees with height  $\ell - 1$ , branching  $B$ , and regularity constant  $C$ . Let  $T$  be an Ahlfors-David regular tree with height  $\ell$ , branching  $B$ , and regularity constant  $C$ , and let  $T'$  be a pruned subtree of  $T$ . Then at least one of the vertices in  $\mathcal{V}_1(T)$  is missing from  $T'$ . We call this vertex  $v^*$ .

By Lemma 6.17, each of the trees  $\{T_v \mid v \in \mathcal{V}_1(T')\}$  is a regular tree with height  $\ell - 1$ , branching  $B$ , and regularity constant  $C$ . Thus we can apply the induction hypothesis to each such tree to obtain

$$\begin{aligned} |\mathcal{L}(T')| &= \sum_{v \in \mathcal{V}_1(T')} |\mathcal{L}(T'_v)| \\ &\leq (1 - C^{-2}B^{-1})^{\ell-1} \sum_{v \in \mathcal{V}_1(T')} |\mathcal{L}(T_v)| \\ &\leq (1 - C^{-2}B^{-1})^{\ell-1} (|\mathcal{L}(T)| - |\mathcal{L}(T_{v^*})|). \end{aligned}$$

By (6.28) we have

$$|\mathcal{L}(T_{v^*})| \geq C^{-2}B^{-1}|\mathcal{L}(T)|. \quad (6.38)$$

Thus

$$|\mathcal{L}(T)| - |\mathcal{L}(T_{v^*})| \leq (1 - C^{-2}B^{-1}) \cdot |\mathcal{L}(T)|,$$

so

$$|\mathcal{L}(T')| \leq (1 - C^{-2}B^{-1})^\ell |\mathcal{L}(T)|. \quad \square$$

**6.7. Finishing the proof of Theorem 6.** Let  $\mathcal{X}$  be a  $\delta$ -regular set with regularity constant  $C$  and let  $\alpha > 0$ . Let

$$M = \lceil M_0 \rceil, \quad N = \left\lfloor \frac{\log(1/\alpha)}{\log M} \right\rfloor,$$

where  $M_0$  is defined in (6.33), and let  $T_{\mathcal{X};M,N}$  be the associated Ahlfors-David regular tree constructed in §6.4.

Let  $T'$  be the subtree of  $T_{\mathcal{X};M,N}^3$  which consists of triples of intervals that hit  $\mathcal{X}$  (see §6.5). By Proposition 6.20,  $T'$  is pruned in the sense of §6.6. By Lemmas 6.17 and 6.22 and the inequality  $(1-t) \leq e^{-t}$ , we have

$$\begin{aligned} |\mathcal{L}(T')| &\leq \left(1 - (C_2^6 M^{3\delta})^{-1}\right)^N C_2^3 M^{3\delta N} \\ &\leq C_2^3 M^{N(3\delta-\rho)}, \quad \rho = (C_2^6 M^{3\delta} \log M)^{-1}. \end{aligned} \quad (6.39)$$

Expanding the definition of  $C_2$  and  $M_0$  from (6.30) and (6.33), we have

$$\rho \geq \delta \exp \left[ -K_6(1-\delta)^{-14}(1 + \log^{14} C) \right],$$

where  $K_6$  is a large absolute constant.

Now, assume that  $(x_1, x_2, x_3) \in \mathcal{X}^3$  satisfies  $x_1 - x_2 + x_3 \in \mathcal{X}(\alpha)$ . Then  $x_i \in I_i$ ,  $i = 1, 2, 3$ , where  $I_i$  are intervals corresponding to some leaves  $v_i$  of  $T_{\mathcal{X};M,N}$ . Moreover, since  $\alpha \leq M^{-N}$ , the vertex  $(v_1, v_2, v_3)$  hits  $\mathcal{X}$  and thus is a leaf of  $T'$ . By (1.23), for each leaf  $(v_1, v_2, v_3)$  of  $T'$  we have

$$\mu_\delta^4(\{(x_1, x_2, x_3, x_4) \in \mathcal{X}^4 \mid |x_1 - x_2 + x_3 - x_4| \leq \alpha, x_i \in I_{v_i}, i = 1, 2, 3\}) \leq C_4 \alpha^{4\delta}$$

for some constant  $C_4$  depending on  $C$  and  $\delta$  but not on  $\alpha$ . Combining this with (6.39) and recalling (6.1), we conclude that

$$\mathcal{E}_A(\mathcal{X}, \mu_\delta, \alpha) \leq C_4 C_2^3 \alpha^{\delta+\rho} \leq C_4 C_2^3 \alpha^{\delta+\beta_{\mathcal{X}}}$$

where  $\beta_{\mathcal{X}} = \delta \exp \left[ -\mathbf{K}(1-\delta)^{-14}(1 + \log^{14} C) \right]$  for  $\mathbf{K}$  a large absolute constant. This finishes the proof of Theorem 6.

## 6.8. Further remarks.

6.8.1. *A discretized additive energy bound.* We have chosen to phrase Theorem 6 in the language of Ahlfors-David regular sets. However, we only examine these sets at scales between  $\alpha$  and 1. Our proof of Theorem 6 also gives the following variant:

**Proposition 6.23.** *Let  $\mathcal{X} \subset [0, 1]$  be a union of intervals of length  $\alpha$ . Suppose that for all  $\alpha \leq r \leq 1$  and all  $x \in \mathcal{X}$  we have the bounds*

$$C_{\mathcal{X}}^{-1} r^\delta \alpha^{1-\delta} \leq \mu_L(\mathcal{X} \cap B(x, r)) \leq C_{\mathcal{X}} r^\delta \alpha^{1-\delta}, \quad (6.40)$$

where  $\mu_L$  is the one-dimensional Lebesgue measure. Then

$$\mu_L^4(\{(x_1, x_2, x_3, x_4) \in \mathcal{X} \mid |x_1 - x_2 + x_3 - x_4| < \alpha\}) \leq \tilde{C} \alpha^{4-3\delta+\beta_{\mathcal{X}}}, \quad (6.41)$$

where  $\beta_{\mathcal{X}}$  is as given in (6.3) and  $\tilde{C}$  is some constant.

6.8.2. *Higher dimensions.* Most of the arguments in this section extend to higher dimensions without difficulty. The real issue is extending Theorem 6 to  $\delta \geq 1$ . This may be challenging because Proposition 6.13 may not be true if  $\delta \geq 1$ . For example, a unit line segment in  $[0, 1]^2$  is a 1-regular set, but it contains arbitrarily long arithmetic progressions. A unit line segment in  $[0, 1]^2$  also has maximal additive energy, so no variant of Theorem 6 can hope to hold for that example. If  $\delta \geq 1$ , it is not clear what the proper hypotheses for the theorem should be.

One possible avenue is the following. In [BLZ], the authors study 1-regular sets that satisfy an additional property called  $(\rho, C_1)$ -unrectifiability. If a 1-regular set is  $(\rho, C_1)$ -unrectifiable, then for all rectangles  $R$  of dimensions  $r_1 \geq r_2$ , we have  $\mu_1(\mathcal{X} \cap R) \leq Cr_1^{1-\rho}r_2^\rho$ . If  $\rho > 0$  then sets with this property strongly avoid arithmetic progressions. It is possible that this property can be generalized to other  $\delta$ -regular sets for  $\delta > 1$ .

6.8.3. *Improving the bounds on  $\beta_{\mathcal{X}}$ .* It is likely that the bound on  $\beta_{\mathcal{X}}$  from Theorem 6 can be substantially improved. A modest improvement would be to replace the bound in (6.3) by  $C_{\mathcal{X}}^{-K/(1-\delta)^K}$  where  $K$  is an absolute constant.

However, the following example shows that the constant  $\beta_{\mathcal{X}}$  must go to 0 as  $C \rightarrow \infty$ . Let  $C > 1$  be an integer and let

$$\mathcal{X} = \{x \in [0, 1] \mid x \text{ has a base } C \text{ expansion of the form } x = 0.a_10a_30a_5 \dots\}. \quad (6.42)$$

Then  $\mathcal{X}$  is a  $1/2$ -regular set with regularity constant  $C$ . On the other hand,

$$\mathcal{E}_A(\mathcal{X}, \alpha) \geq \alpha^{1/2 + \frac{1}{10 \log C}}.$$

A similar example can be constructed for other values of  $\delta$ . This shows that in Theorem 6 we cannot take  $\beta_{\mathcal{X}} > 0$  to be independent of  $C$ .

## 7. REGULARITY AND ADDITIVE ENERGY OF LIMIT SETS

In this section, we study regularity of limit sets of convex co-compact groups and their stereographic projections (§7.1). We next use the results of §6 to prove Theorem 5 (§7.2).

7.1. **Regularity of limit sets.** In this section, we consider a convex co-compact hyperbolic quotient  $M = \Gamma \backslash \mathbb{H}^n$  and use the Ahlfors-David regularity of the limit set  $\Lambda_\Gamma \subset \mathbb{S}^{n-1}$  to establish Ahlfors-David regularity of the stereographic projections  $\mathcal{G}(y_0, \Lambda_\Gamma) \subset T_{y_0} \mathbb{S}^{n-1}$ ,  $y_0 \in \Lambda_\Gamma$ , where  $\mathcal{G}$  is defined in (1.19):

**Lemma 7.1.** *Let  $\mathbf{C}$  be the constant in Ahlfors-David regularity of the limit set, as defined in (1.24). Then for each  $y_0, y_1 \in \Lambda_\Gamma$ ,  $y_0 \neq y_1$ , we have*

$$\mathbf{K}_0^{-1} \mathbf{C}^{-1} r^\delta \leq \mu_\delta(\mathcal{G}(y_0, \Lambda_\Gamma) \cap B(\mathcal{G}(y_0, y_1), r)) \leq \mathbf{K}_0 \mathbf{C} r^\delta, \quad r > 0 \quad (7.1)$$

where  $\mu_\delta$  is the  $\delta$ -Hausdorff measure on  $T_{y_0}\mathbb{S}^{n-1}$  and  $\mathbf{K}_0$  is a global constant (depending only on the dimension).

The case of bounded  $r$  in the above lemma is an immediate consequence of the regularity of  $\Lambda_\Gamma$ . To show that the constants in the regularity statement do not deteriorate when  $r \rightarrow \infty$ , we will prove that if we shrink the set  $\mathcal{G}(y_0, \Lambda_\Gamma) - \mathcal{G}(y_0, y_1)$ , we obtain an isometric image of the set  $\mathcal{G}(y'_0, \Lambda_\Gamma) - \mathcal{G}(y'_0, y'_1)$  for some other  $y'_0, y'_1 \in \Lambda_\Gamma$ . For that we use the group action, or equivalently, argue on the quotient manifold  $M$  rather than on  $\mathbb{H}^n$ .

We first write the sets  $\mathcal{G}(y_0, \Lambda_\Gamma) \subset T_{y_0}\mathbb{S}^{n-1}$  as subsets of the unstable spaces on  $M$  via a map  $\mathcal{U}_-$  constructed using horocyclic flows (see Appendix A for the proof):

**Lemma 7.2.** *Let  $S^*\mathbb{H}^n \subset T^*\mathbb{H}^n$  be the unit cotangent bundle and  $E_u$  the unstable foliation, see (4.3). Then there exists a smooth map*

$$\mathcal{U}_- : \{(x, \xi, \eta) \mid (x, \xi) \in S^*\mathbb{H}^n, \eta \in E_u(x, \xi)\} \rightarrow S^*\mathbb{H}^n$$

such that for  $(\tilde{x}, \tilde{\xi}) := \mathcal{U}_-(x, \xi, \eta)$  and  $B_\pm : S^*\mathbb{H}^n \rightarrow \mathbb{S}^{n-1}$  defined in (4.7),

$$B_-(\tilde{x}, \tilde{\xi}) = B_-(x, \xi), \quad \mathcal{P}(\tilde{x}, B_-(x, \xi)) = \mathcal{P}(x, B_-(x, \xi)), \quad (7.2)$$

$$\mathcal{G}(B_-(x, \xi), B_+(\tilde{x}, \tilde{\xi})) - \mathcal{G}(B_-(x, \xi), B_+(x, \xi)) = \mathcal{P}(x, B_-(x, \xi))\mathcal{T}_-(x, \xi)\eta, \quad (7.3)$$

where  $\mathcal{T}_-(x, \xi) : E_u(x, \xi) \rightarrow T_{B_-(x, \xi)}\mathbb{S}^{n-1}$  is some linear isometry and  $\mathcal{P}$  is the Poisson kernel defined in (4.43). Moreover, the map  $\mathcal{U}_-$  commutes with the natural action of the isometry group  $\text{PSO}(1, n)$  and thus descends to a map

$$\mathcal{U}_- : \{(x, \xi, \eta) \mid (x, \xi) \in S^*M, \eta \in E_u(x, \xi)\} \rightarrow S^*M.$$

**Remarks.** (i) For each  $(x, \xi) \in S^*M$ , the set  $\{\mathcal{U}_-(x, \xi, \eta) \mid \eta \in E_u(x, \xi)\}$  is the unstable manifold passing through  $(x, \xi)$ , and the differential of the map  $\eta \mapsto \mathcal{U}_-(x, \xi, \eta)$  at  $\eta = 0$  is the embedding  $E_u(x, \xi) \rightarrow T_{(x, \xi)}(S^*M)$ . See Figure 9. The map  $\mathcal{T}_-$  is related to the parametrization of  $E_u(x, \xi)$  by the orthogonal complement  $\mathcal{E}(x, \xi) \subset T_xM$  of  $\xi$  (see the paragraph preceding (4.5)) and to the parallel transport map  $\mathcal{E}(x, \xi) \rightarrow T_{B_-(x, \xi)}\mathbb{S}^{n-1}$  (see [DFG, §3.6]), but we do not need an explicit expression for  $\mathcal{T}_-$  here. (ii) In dimension 2,  $\mathcal{U}_-$  is given by the flow of the unstable horocyclic vector field  $U_-$  (see for instance [DFG, (2.1)]):

$$\mathcal{U}_-(x, \xi, \eta) = e^{sU_-}(x, \xi), \quad \eta = sU_-(x, \xi) \in E_u(x, \xi), \quad s \in \mathbb{R}.$$

For each point  $(x, \xi) \in K \cap S^*M$ , where  $K = \Gamma_+ \cap \Gamma_-$  is the trapped set, define

$$\mathcal{F}_{(x, \xi)} := \{\eta \mid \mathcal{U}_-(x, \xi, \eta) \in K\} \subset E_u(x, \xi). \quad (7.4)$$

Note that  $(\tilde{x}, \tilde{\xi}) := \mathcal{U}_-(x, \xi, \eta)$  lies in  $\Gamma_+$  for all  $\eta$ , since the geodesics starting at  $(x, \xi)$  and  $(\tilde{x}, \tilde{\xi})$  converge to each other as  $t \rightarrow -\infty$  by (7.2) (see also (4.12)); therefore,  $K$

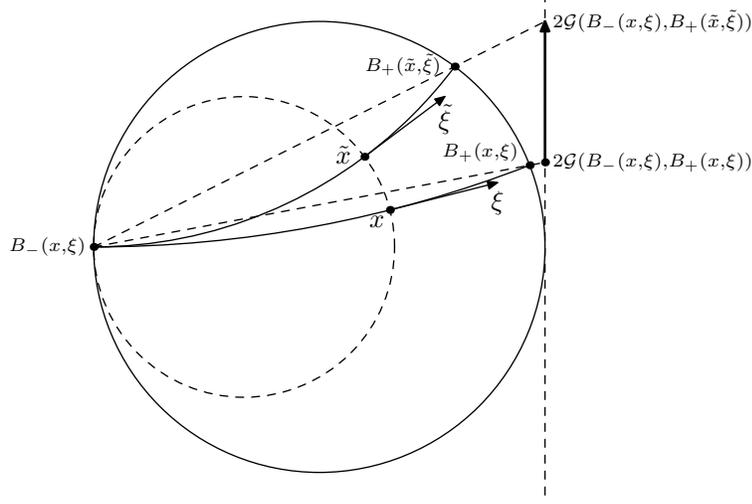


FIGURE 9. The points  $(x, \xi) \in S^*\mathbb{H}^n$  and  $(\tilde{x}, \tilde{\xi}) = \mathcal{U}_-(x, \xi, \eta)$ . The dashed circle is a horocycle and the solid arcs are geodesics through  $(x, \xi)$  and  $(\tilde{x}, \tilde{\xi})$ .

can be replaced in (7.4) by  $\Gamma_-$ . By (7.3) and (4.12), for each  $(x, \xi) \in S^*\mathbb{H}^n$  we have

$$\mathcal{G}(y_-, \Lambda_\Gamma) - \mathcal{G}(y_-, y_+) = \mathcal{P}(x, y_-) \mathcal{T}_-(x, \xi) \mathcal{F}_{\pi_\Gamma(x, \xi)}, \quad y_\pm := B_\pm(x, \xi), \quad (7.5)$$

with  $\pi_\Gamma$  defined in (4.4).

*Proof of Lemma 7.1.* Let  $y_0, y_1 \in \Lambda_\Gamma$  and  $y_0 \neq y_1$ .

We first prove (7.1) for the case  $0 < r < 1$ . Define the diffeomorphism

$$\Phi : \mathbb{S}^{n-1} \setminus \{y_0\} \rightarrow T_{y_0} \mathbb{S}^{n-1}, \quad \Phi(y) := \mathcal{G}(y_0, y).$$

Then  $d\Phi(y)$  is conformal with factor  $\frac{1}{2}(1 + |\Phi(y)|^2)$ . It follows that

$$\sup_{\Phi(y) \in B(\eta_1, 1)} \|d\Phi(y)\| \leq \frac{3}{2}(1 + |\eta_1|^2), \quad \sup_{\Phi(y) \in B(\eta_1, 1)} \|d\Phi(y)^{-1}\| \leq \frac{6}{1 + |\eta_1|^2}$$

where  $\eta_1 := \Phi(y_1)$ , and therefore

$$B\left(y_1, \frac{2r}{3(1 + |\eta_1|^2)}\right) \subset \Phi^{-1}(B(\eta_1, r)) \subset B\left(y_1, \frac{6r}{1 + |\eta_1|^2}\right).$$

We now have by (1.23)

$$\begin{aligned} \mu_\delta(\Phi(\Lambda_\Gamma) \cap B(\eta_1, r)) &\leq \left(\frac{3}{2}(1 + |\eta_1|^2)\right)^\delta \mu_\delta(\Lambda_\Gamma \cap \Phi^{-1}(B(\eta_1, r))) \\ &\leq \left(\frac{3}{2}(1 + |\eta_1|^2)\right)^\delta \mu_\delta\left(\Lambda_\Gamma \cap B\left(y_1, \frac{6r}{1 + |\eta_1|^2}\right)\right) \\ &\leq \mathbf{C}(9r)^\delta \end{aligned}$$

and similarly

$$\mu_\delta(\Phi(\Lambda_\Gamma) \cap B(\eta_1, r)) \geq \left(\frac{1 + |\eta_1|^2}{6}\right)^\delta \mu_\delta(\Lambda_\Gamma \cap \Phi^{-1}(B(\eta_1, r))) \geq \mathbf{C}^{-1}(r/9)^\delta$$

which gives (7.1) for  $0 < r < 1$  with  $\mathbf{K}_0 := 9^{n-1} \geq 9^\delta$ .

Now, assume that  $r \geq 1$ . Take some  $(x, \xi) \in S^*\mathbb{H}^n$  on the geodesic connecting  $y_0$  and  $y_1$ , that is

$$B_-(x, \xi) = y_0, \quad B_+(x, \xi) = y_1.$$

Let  $\gamma \in \Gamma$  and put

$$(x', \xi') := \gamma \cdot (x, \xi), \quad B_-(x', \xi') = y'_0, \quad B_+(x', \xi') = y'_1;$$

note that  $y'_0, y'_1 \in \Lambda_\Gamma$ . We choose  $(x, \xi)$  and  $\gamma$  such that

$$r' := \frac{\mathcal{P}(x', y'_0)}{\mathcal{P}(x, y_0)} r < 1. \quad (7.6)$$

To do that, we first remark that there exists  $R > 0$  depending on the quotient  $M$  such that for each  $(\tilde{x}, \tilde{\xi}) \in K \cap S^*M$ , there exists

$$(x', \xi') \in \pi_\Gamma^{-1}(\tilde{x}, \tilde{\xi}), \quad \mathcal{P}(x', B_-(x', \xi')) \leq R. \quad (7.7)$$

This follows immediately from the compactness of  $K \cap S^*M$ . To ensure (7.6), it remains to take  $(x, \xi)$  such that  $\mathcal{P}(x, y_0) > rR$  (which is always possible since the function  $\mathcal{P}(x, y_0)$  grows exponentially along the backwards geodesic flow) and choose  $(x', \xi')$  using (7.7) with  $(\tilde{x}, \tilde{\xi}) := \pi_\Gamma(x, \xi)$ .

From (7.5) and the fact that  $\pi_\Gamma(x, \xi) = \pi_\Gamma(x', \xi')$ , we have

$$\mathcal{G}(y_0, \Lambda_\Gamma) - \mathcal{G}(y_0, y_1) = \frac{\mathcal{P}(x, y_0)}{\mathcal{P}(x', y'_0)} \tilde{\mathcal{T}}(\mathcal{G}(y'_0, \Lambda_\Gamma) - \mathcal{G}(y'_0, y'_1)) \quad (7.8)$$

where  $\tilde{\mathcal{T}} : T_{y'_0}\mathbb{S}^{n-1} \rightarrow T_{y_0}\mathbb{S}^{n-1}$  is an isometry. Since (7.1) is already known for  $r < 1$ , we have using (7.6),

$$\mathbf{K}_0^{-1} \mathbf{C}^{-1}(r')^\delta \leq \mu_\delta(\mathcal{G}(y'_0, \Lambda_\Gamma) \cap B(\mathcal{G}(y'_0, y'_1), r')) \leq \mathbf{K}_0 \mathbf{C}(r')^\delta.$$

Combining this with (7.8), we obtain (7.1).  $\square$

**7.2. Regularity, additive energy, and Minkowski dimension.** In this section, we state a few results estimating the Lebesgue measure of neighborhoods of Ahlfors-David regular sets. This establishes bounds on Minkowski dimensions of these sets. We rely on the following

**Definition 7.3.** *Assume that  $(\mathcal{M}, d)$  is a metric space. For  $\mathcal{X} \subset \mathcal{M}$  and  $\alpha > 0$ , define the maximal number of  $\alpha$ -separated points in  $\mathcal{X}$ :*

$$\mathcal{N}(\mathcal{X}, \alpha) = \max\{N \mid x_1, \dots, x_N \in \mathcal{X}, d(x_i, x_j) > \alpha \text{ for } i \neq j\}.$$

For regular sets, the quantity  $\mathcal{N}(\mathcal{X}, \alpha)$  establishes a link between the Hausdorff and Minkowski dimensions:

**Lemma 7.4.** *Let  $(\mathcal{M}, d)$  be a metric space and let  $\mathcal{X} \subset \mathcal{M}$  be compact.*

1. *If  $\mathcal{X}$  is  $\delta$ -regular in the sense of Definition 1.4 with constant  $C_{\mathcal{X}}$ , then for each  $x \in \mathcal{X}$ ,*

$$C_{\mathcal{X}}^{-2} \left( \frac{\alpha'}{\alpha} \right)^{\delta} \leq \mathcal{N}(\mathcal{X} \cap B(x, \alpha'), \alpha) \leq C_{\mathcal{X}}^2 \left( 1 + \frac{2\alpha'}{\alpha} \right)^{\delta}, \quad 0 < \alpha, \alpha' < \text{diam}(\mathcal{M}).$$

2. *If  $\mathcal{M}$  is an  $m$ -dimensional Riemannian manifold and  $\mathcal{X}(\alpha)$  is the  $\alpha$ -neighborhood of  $\mathcal{X}$ , then*

$$C^{-1} \alpha^m \mathcal{N}(\mathcal{X}, \alpha) \leq \mu_L(\mathcal{X}(\alpha)) \leq C \alpha^m \mathcal{N}(\mathcal{X}, \alpha), \quad 0 < \alpha < 1,$$

where  $\mu_L$  is the Lebesgue measure induced by the metric and  $C$  is some constant independent of  $\alpha$ .

*Proof.* We will begin with the first statement. Fix a ball  $B(x, \alpha')$  with  $x \in \mathcal{X}$ . Let  $\{x_1, \dots, x_N\} \subset \mathcal{X} \cap B(x, \alpha')$  be a maximal collection of  $\alpha$ -separated points. Then

$$\mathcal{X} \cap B(x, \alpha') \subset \bigcup_{i=1}^N B(x_i, \alpha), \quad \bigsqcup_{i=1}^N B\left(x_i, \frac{\alpha}{2}\right) \subset B\left(x, \alpha' + \frac{\alpha}{2}\right).$$

It follows that

$$\begin{aligned} N &\geq C_{\mathcal{X}}^{-1} \alpha^{-\delta} \sum_{i=1}^N \mu_{\delta}(\mathcal{X} \cap B(x_i, \alpha)) \geq C_{\mathcal{X}}^{-1} \alpha^{-\delta} \mu_{\delta}(\mathcal{X} \cap B(x, \alpha')) \\ &\geq C_{\mathcal{X}}^{-2} \left( \frac{\alpha'}{\alpha} \right)^{\delta}, \\ N &\leq \left( \frac{2}{\alpha} \right)^{\delta} C_{\mathcal{X}} \sum_{i=1}^N \mu_{\delta}\left(\mathcal{X} \cap B\left(x_i, \frac{\alpha}{2}\right)\right) \leq \left( \frac{2}{\alpha} \right)^{\delta} C_{\mathcal{X}} \mu_{\delta}\left(\mathcal{X} \cap B\left(x, \alpha' + \frac{\alpha}{2}\right)\right) \\ &\leq C_{\mathcal{X}}^2 \left( 1 + \frac{2\alpha'}{\alpha} \right)^{\delta}; \end{aligned}$$

this finishes the proof of the first statement.

The second statement follows similarly from the inclusions

$$\mathcal{X}(\alpha) \subset \bigcup_{i=1}^N B(x_i, 2\alpha), \quad \bigsqcup_{i=1}^N B\left(x_i, \frac{\alpha}{2}\right) \subset \mathcal{X}(\alpha),$$

where  $\{x_1, \dots, x_N\} \subset \mathcal{X}$  is a maximal collection of  $\alpha$ -separated points, plus the observation that on any compact subset  $\Omega \subset \mathcal{M}$ , the Lebesgue measure of a ball of radius  $\alpha$  is between  $C_{\Omega}^{-1} \alpha^m$  and  $C_{\Omega} \alpha^m$ .  $\square$

As an application, we obtain

*Proof of (5.3).* Follows directly from  $\delta$ -regularity of the limit set (see §1.3), Lemma 7.4, and the fact that  $\Lambda_\Gamma(\alpha) \cap B(y_0, \alpha')$  contains the  $\alpha$ -neighborhood of  $\Lambda_\Gamma \cap B(y_0, \alpha' - \alpha)$  and is contained in the  $\alpha$ -neighborhood of  $\Lambda_\Gamma \cap B(y_0, \alpha' + \alpha)$ .  $\square$

**Remark.** In Definition 1.4 of Ahlfors-David regularity we used the Hausdorff measure. However, any other Borel measure could be used instead:

**Lemma 7.5** ([DaSe], Lemma 1.2). *Let  $(\mathcal{M}, d)$  be a complete metric space with more than one element and let  $\mathcal{X} \subset \mathcal{M}$ . Let  $\mu$  be a Borel measure on  $\mathcal{M}$  with the property that for all  $x \in \mathcal{X}$ ,*

$$C_{\mathcal{X}}^{-1}r^\delta \leq \mu(\mathcal{X} \cap B(x, r)) \leq C_{\mathcal{X}}r^\delta, \quad 0 < r < \text{diam}(\mathcal{M}). \quad (7.9)$$

*Then  $\mathcal{X}$  is  $\delta$ -regular. The regularity constant depends only on  $\delta$  and  $C_{\mathcal{X}}$ .*

We now give the proof of Theorem 5. The first step is the following

**Lemma 7.6.** *Let  $\Lambda_\Gamma$  be the limit set of a convex co-compact hyperbolic surface,  $\mathbf{C}$  be defined in (1.24),  $\mathbf{K}_0$  be given in Lemma 7.1, and  $\mathcal{G}$  be defined in (1.19). Take  $y_0 \in \Lambda_\Gamma$  and  $R > 0$ . Then there exists an interval  $[-1, 1] \subset I \subset [-2, 2]$  such that*

$$\mathcal{X} := I \cap R^{-1}\mathcal{G}(y_0, \Lambda_\Gamma) \subset T_{y_0}\mathbb{S}^1 \simeq \mathbb{R}$$

*is  $\delta$ -regular with regularity constant  $\mathbf{C}_2 := (50\mathbf{K}_0\mathbf{C})^{\frac{1+\delta}{1-\delta}}$ .*

*Proof.* By Lemma 7.1, the set

$$\mathcal{Y} := R^{-1}\mathcal{G}(y_0, \Lambda_\Gamma) \subset \mathbb{R}$$

is  $\delta$ -regular with constant  $\mathbf{K}_0\mathbf{C}$ . Divide  $[-2, -1]$  and  $[1, 2]$  into  $\mathbf{C}_1$  intervals of size  $\mathbf{C}_1^{-1}$  each, where  $\mathbf{C}_1 := \lceil (10\mathbf{K}_0^2\mathbf{C}^2)^{\frac{1}{1-\delta}} \rceil$ . By Proposition 6.13, at least one of the sub-intervals in  $[-2, -1]$  and at least one of the sub-intervals in  $[1, 2]$  must be disjoint from  $\mathcal{Y}$ . Call these intervals  $I_1$  and  $I_2$ . Let  $I$  be the convex hull of the midpoints of  $I_1$  and  $I_2$ .

We will show that  $\mathcal{X} := I \cap \mathcal{Y}$  is  $\delta$ -regular. Let  $x \in \mathcal{X}$  and  $0 < r \leq 4$ . We immediately have

$$\mu_\delta(\mathcal{X} \cap B(x, r)) \leq \mu_\delta(\mathcal{Y} \cap B(x, r)) \leq \mathbf{K}_0\mathbf{C}r^\delta.$$

It remains to prove a lower bound. If  $r < \mathbf{C}_1^{-1}$  then  $\mathcal{Y} \cap B(x, r) \subset I$  and thus

$$\mu_\delta(\mathcal{X} \cap B(x, r)) = \mu_\delta(\mathcal{Y} \cap B(x, r)) \geq (\mathbf{K}_0\mathbf{C})^{-1}r^\delta.$$

On the other hand, if  $\mathbf{C}_1^{-1} \leq r \leq 4$ , then

$$\begin{aligned} \mu_\delta(\mathcal{X} \cap B(x, r)) &\geq \mu_\delta(\mathcal{X} \cap B(x, \mathbf{C}_1^{-1})) \\ &\geq \mathbf{K}_0^{-1}\mathbf{C}^{-1}\mathbf{C}_1^{-\delta} \\ &\geq \mathbf{C}_2^{-1}r^\delta. \end{aligned} \quad \square$$

We finally combine Theorem 6 from §6, Lemma 7.4, and Lemma 7.6 to obtain

*Proof of Theorem 5.* Let  $\Lambda_\Gamma$  be the limit set of a convex co-compact hyperbolic group. Let  $y_0 \in \Lambda_\Gamma$ . Let  $\alpha > 0$  (small) and  $C_1 \geq 1$  (large), and put  $\alpha_1 := C_1^{-1}\alpha$ . First, note that

$$E_A(\mathcal{G}(y_0, \Lambda_\Gamma) \cap B(0, C_1), \alpha) = E_A(C_1^{-1}\mathcal{G}(y_0, \Lambda_\Gamma) \cap B(0, 1), \alpha_1) \leq E_A(\mathcal{X}, \alpha_1), \quad (7.10)$$

where  $\mathcal{X} := I \cap C_1^{-1}\mathcal{G}(y_0, \Lambda_\Gamma) \subset [-2, 2]$  is defined in Lemma 7.6.

Each point  $(\eta_1, \eta_2, \eta_3, \eta_4) \in \mathcal{X}(\alpha_1)^4$  satisfying  $|\eta_1 - \eta_2 + \eta_3 - \eta_4| \leq \alpha_1$  lies in the  $4\alpha_1$ -neighborhood of the set

$$\mathcal{Z}_{\alpha_1} = \{(\eta_1, \eta_2, \eta_3, \eta_4) \in \mathcal{X}^4 \mid |\eta_1 - \eta_2 + \eta_3 - \eta_4| \leq 5\alpha_1\}.$$

By Definition 1.2 and part 2 of Lemma 7.4, we have for some global constant  $C$ ,

$$E_A(\mathcal{X}, \alpha_1) \leq CN(\mathcal{Z}_{\alpha_1}, 4\alpha_1). \quad (7.11)$$

We next claim that

$$\mathcal{N}(\mathcal{Z}_{\alpha_1}, 4\alpha_1) \leq C\alpha_1^{-4\delta}\mathcal{E}_A(\mathcal{X}, \mu_\delta, 9\alpha_1) \quad (7.12)$$

where  $\mathcal{E}_A$  is given by Definition 6.1 and  $C$  is some constant depending on  $\mathbf{C}$ . Indeed, let  $z_1, \dots, z_N \in \mathcal{Z}_{\alpha_1}$  be a  $4\alpha_1$ -separated set of points. Then

$$\bigsqcup_{j=1}^N B(z_j, 2\alpha_1) \subset \{|\eta_1 - \eta_2 + \eta_3 - \eta_4| \leq 9\alpha_1\}.$$

Taking the  $\mu_\delta^4$  measure of the intersection of both sides with  $\mathcal{X}^4$  and arguing similarly to the proof of part 1 of Lemma 7.4, we obtain (7.12).

Finally, applying Theorem 6 to  $\frac{1}{2}\mathcal{X}$ , which is  $\delta$ -regular by Lemma 7.6, we obtain

$$\mathcal{E}_A(\mathcal{X}, \mu_\delta, 9\alpha_1) \leq C\alpha_1^{\delta+\beta_E}. \quad (7.13)$$

Here  $\beta_E = \delta \exp[-\mathbf{K}(1-\delta)^{-28}(1+\log^{14}\mathbf{C})]$ , where  $\mathbf{K}$  is an absolute constant, and  $C$  is some constant depending on  $\mathbf{C}$ .

Combining (7.10)–(7.13), we conclude that  $\Lambda_\Gamma$  satisfies the additive energy bound with exponent  $\beta_E$  in the sense of Definition 1.3.  $\square$

**7.3. Example: three-funneled surfaces.** We now consider a particular family of convex co-compact hyperbolic surfaces and show that the regularity constants in Lemma 7.1 for the corresponding limit sets have a uniform upper bound when the surface varies in a compact set in the moduli space. (Similar reasoning is expected to work for general convex co-compact hyperbolic surfaces.)

More precisely, we study the family of *three-funneled surfaces*  $M_\ell$ , parametrized by the Fenchel–Nielsen coordinates  $\ell := (\ell_1, \ell_2, \ell_3) \in (0, \infty)^3$ . To construct  $M_\ell$ , we start with a right-angled hyperbolic hexagon with sides  $\ell_1/2, q_3, \ell_2/2, q_1, \ell_3/2, q_2$ . This

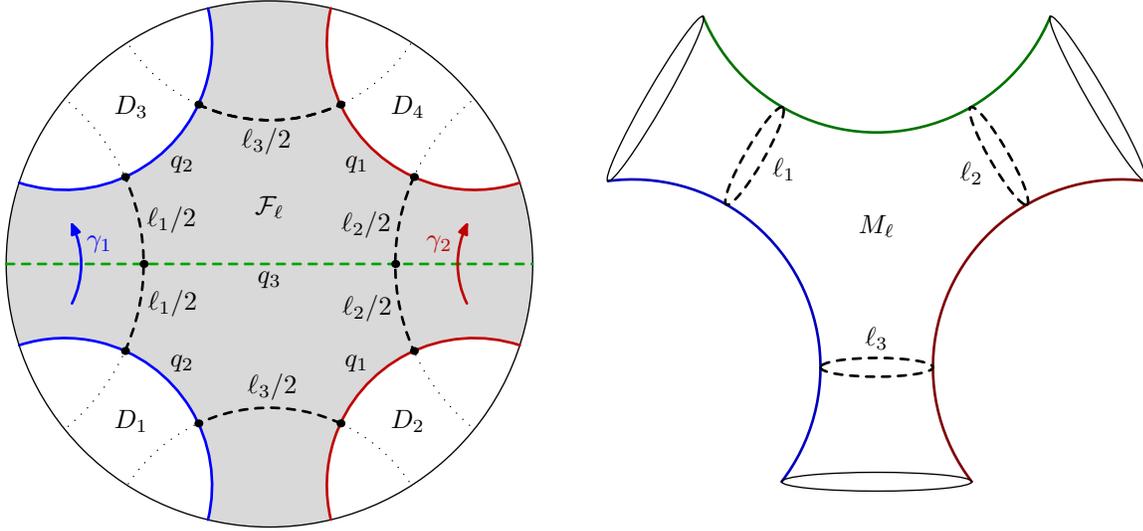


FIGURE 10. A three-funneled surface (on the right) and its fundamental domain in the Poincaré disk model (on the left).

hexagon is unique up to isometry, and  $q_1, q_2, q_3 > 0$  are determined by a formula involving only  $\ell_1, \ell_2$ , and  $\ell_3$  (see for example [Ra, Theorem 3.5.13]). Gluing two such hexagons along the  $q_3$  side, we obtain a right-angled hyperbolic octagon with sides  $\ell_1, q_2, \ell_3/2, q_1, \ell_2, q_1, \ell_3/2, q_2$ .

Attaching funnel ends along the  $\ell_1, \ell_2, \ell_3/2$  sides to the above octagon, we obtain a fundamental domain  $\mathcal{F}_\ell \subset \mathbb{H}^2$ . The complement of  $\mathcal{F}_\ell$  is the disjoint union of four open geodesic half-disks (i.e. regions of  $\mathbb{H}^2$  bounded by a geodesic)  $D_1, D_2, D_3, D_4$ , where  $D_1, D_3$  are bounded by the geodesics containing the two  $q_2$  sides of the octagon and  $D_2, D_4$  are bounded by the geodesics containing its  $q_1$  sides. See Figure 10.

We next define the group  $\Gamma_\ell \subset \text{PSL}(2, \mathbb{R})$  using the following Schottky representation. Let  $\omega_1, \omega_2$  be the geodesics on  $\mathbb{H}^2$  containing the  $\ell_1$  and  $\ell_2$  sides of the octagon. Then there exist unique  $\gamma_1, \gamma_2 \in \text{PSL}(2, \mathbb{R})$  satisfying

$$\gamma_j(\overline{D_j}) = \mathbb{H}^2 \setminus D_{j+2}, \quad \gamma_j(\omega_j) = \omega_j, \quad j = 1, 2. \quad (7.14)$$

Let  $\Gamma_\ell$  be the group generated by  $\gamma_1, \gamma_2$ . Then  $\Gamma_\ell$  is a free group, and the quotient

$$M_\ell := \Gamma_\ell \backslash \mathbb{H}^2$$

is a convex co-compact hyperbolic surface with fundamental domain  $\mathcal{F}_\ell$ . The numbers  $\ell_1, \ell_2, \ell_3$  are the lengths of the geodesic necks separating the funnels of  $M_\ell$  from the convex core. See for instance [Bo07, §15.1] for details. The octagon constructed above, the disks  $D_j$ , and the group elements  $\gamma_j$  depend continuously on the choice of  $\ell = (\ell_1, \ell_2, \ell_3)$ .

Denote by  $\Lambda_\ell = \Lambda_{\Gamma_\ell}$  the limit set of  $\Gamma_\ell$  and by  $\delta_\ell$  its dimension. The following proposition states that  $\Lambda_\ell$  is Ahlfors–David regular with regularity constant locally uniform in  $\ell$ ; we use same notation as in Lemma 7.1.

**Proposition 7.7.** *Let  $\mathcal{K} \subset (0, \infty)^3$  be a compact set. Then there exists a constant  $C_{\mathcal{K}} > 0$  such that for each  $\ell \in \mathcal{K}$  and each  $y_0, y_1 \in \Lambda_\ell$ ,  $y_0 \neq y_1$  we have*

$$C_{\mathcal{K}}^{-1} c_\ell r^\delta \leq \mu_\delta(\mathcal{G}(y_0, \Lambda_\ell) \cap B(\mathcal{G}(y_0, y_1), r)) \leq C_{\mathcal{K}} c_\ell r^\delta, \quad r > 0 \quad (7.15)$$

where  $c_\ell > 0$  depends only on  $\ell$ .

Before proving Proposition 7.7, we use it to show the following

**Theorem 7.** *There exists an open set  $\mathcal{U} \subset (0, \infty)^3$  such that:*

- if  $\ell \in (0, \infty)^3$  and  $\delta_\ell = 1/2$ , then  $\ell \in \mathcal{U}$ ;
- if  $\ell \in \mathcal{U}$ , then  $M_\ell$  has an essential spectral gap in the sense of (1.3) of size

$$\beta = \beta_\ell > \max\left(0, \frac{1}{2} - \delta_\ell\right).$$

**Remark.** It is well-known that  $\delta_\ell$  depends continuously on  $\ell$  — see for example [AnRo]. Moreover, there exist  $\ell_-, \ell_+ \in (0, \infty)^3$  such that  $\delta_{\ell_-} < 1/2 < \delta_{\ell_+}$ . In fact, for the case  $\ell_1 = \ell_2 = \ell_3$ , we have  $\delta_\ell \rightarrow 1$  as  $\ell_j \rightarrow 0$  and  $\delta_\ell \rightarrow 0$  as  $\ell_j \rightarrow \infty$  — see [Mc, Theorem 3.5]. By considering a path connecting  $\ell_-$  with  $\ell_+$  and applying Theorem 7, we see that there exist  $\ell$  such that  $\delta_\ell > 1/2$ , yet  $M_\ell$  has an essential spectral gap of size  $\beta_\ell > 0$ .

*Proof.* It suffices to show that for each  $\tilde{\ell}$  with  $\delta_{\tilde{\ell}} = 1/2$ , there exists  $\tilde{\beta} > 0$  and a neighborhood  $U_{\tilde{\ell}}$  of  $\tilde{\ell}$  such that for each  $\ell \in U_{\tilde{\ell}}$ ,  $M_\ell$  has an essential spectral gap of size  $\tilde{\beta}$ . Indeed, it follows from here that there is an open neighborhood  $U'_{\tilde{\ell}}$  of  $\tilde{\ell}$  such that for each  $\ell \in U'_{\tilde{\ell}}$ ,  $M_\ell$  has an essential gap of size  $\beta_\ell > \max(0, 1/2 - \delta_\ell)$ . It remains to let  $\mathcal{U}$  be the union of all  $U'_{\tilde{\ell}}$ .

To show the existence of the neighborhood  $U_{\tilde{\ell}}$ , by Theorems 3 and 4 it suffices to show that there exists a constant  $\beta_E > 0$  such that for all  $\ell$  sufficiently close to  $\tilde{\ell}$ , the set  $\Lambda_\ell$  satisfies the additive energy bound with exponent  $\beta_E$  in the sense of Definition 1.3. To show this, we argue as in the proof of Theorem 5 in §7.2. The only difference is that Lemma 7.1 is replaced by Proposition 7.7. The constant  $c_\ell$  in (7.15) can be removed by Lemma 7.5; alternatively, we may argue using the measure  $c_\ell^{-1} \mu_\delta$  instead of  $\mu_\delta$  since the proof of Theorem 6 never used that  $\mu_\delta$  is the Hausdorff measure.  $\square$

We now prove Proposition 7.7. Assume that  $\ell$  varies in a compact subset of  $(0, \infty)^3$ ; the constants below will depend on that subset. We start with the following

**Lemma 7.8.** *Assume that*

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\ell \subset \mathrm{PSL}(2, \mathbb{R}), \quad |\gamma| := a^2 + b^2 + c^2 + d^2.$$

Then for each  $A \subset \mathbb{S}^1$ , we have

$$\mu_{\delta_\ell}(\Lambda_\ell \cap \gamma(A)) \leq (2|\gamma|)^{\delta_\ell} \mu_{\delta_\ell}(\Lambda_\ell \cap A). \quad (7.16)$$

*Proof.* We identify the upper half-plane model  $\{\text{Im } z > 0\}$  with the disk model  $\{|w| < 1\}$  by the Möbius transformation

$$w = \frac{z - i}{z + i}, \quad z = i \frac{1 + w}{1 - w}.$$

With the Möbius transformation  $\gamma$  given in the  $z$  variable by  $\gamma.z = \frac{az+b}{cz+d}$ , its derivative in the  $w$  variable on  $\mathbb{S}^1$  satisfies

$$|\gamma'(w)| = \frac{1 + z^2}{(az + b)^2 + (cz + d)^2}, \quad w \in \mathbb{S}^1.$$

It follows that  $|\gamma'(w)|^{-1} \leq 2|\gamma|$ ; substituting  $\gamma^{-1}$  instead of  $\gamma$ , we get  $|\gamma'(w)| \leq 2|\gamma|$ . The estimate (7.16) follows from here and the fact that  $\Lambda_\ell \cap \gamma(A) = \gamma(\Lambda_\ell \cap A)$ .  $\square$

Using Lemma 7.8, we next prove

**Lemma 7.9.** *There exists a constant  $c > 0$  such that*

$$\mu_{\delta_\ell}(\Lambda_\ell \cap \overline{D_j}) \geq c\mu_{\delta_\ell}(\Lambda_\ell), \quad j = 1, 2, 3, 4. \quad (7.17)$$

*Proof.* We consider the case  $j = 1$ , the other cases are treated similarly.

By (7.14), we have  $\overline{D_1 \cup D_2 \cup D_4} \subset \gamma_1(\overline{D_1})$ . Since  $|\gamma_1|$  is bounded by some constant depending on  $\mathcal{K}$ , by Lemma 7.8 we obtain for some constant  $C$ ,

$$\mu_{\delta_\ell}(\Lambda_\ell \cap (\overline{D_1 \cup D_2 \cup D_4})) \leq C\mu_{\delta_\ell}(\Lambda_\ell \cap \overline{D_1}). \quad (7.18)$$

Next,  $\overline{D_3} \subset \gamma_2(\overline{D_2})$ . Therefore, similarly to (7.18) we get

$$\mu_{\delta_\ell}(\Lambda_\ell \cap \overline{D_3}) \leq C\mu_{\delta_\ell}(\Lambda_\ell \cap \overline{D_2}). \quad (7.19)$$

Combining (7.18) and (7.19) and using that  $\Lambda_\ell \subset \overline{D_1 \cup D_2 \cup D_3 \cup D_4}$ , we obtain (7.17).  $\square$

We are now ready to give

*Proof of Lemma 7.7.* We argue similarly to the proof of Lemma 7.1. Take  $y_0, y_1 \in \Lambda_\ell$ ,  $y_0 \neq y_1$ , and  $r > 0$ . Fix a large constant  $C_1 > 0$  depending only on  $\mathcal{K}$ , to be chosen later. Let  $(x, \xi) \in S^*\mathbb{H}^2$  be the unique point satisfying

$$B_-(x, \xi) = y_0, \quad B_+(x, \xi) = y_1, \quad \mathcal{P}(x, y_0) = r/C_1.$$

By (4.12), the projection  $\pi_\Gamma(x, \xi) \in S^*M$  lies in the trapped set. Take  $\gamma \in \Gamma_\ell$  such that for  $(x', \xi') := \gamma.(x, \xi)$ , the point  $x'$  lies in the fundamental domain  $\mathcal{F}_\ell$ , and denote

$$y'_0 := B_-(x', \xi'), \quad y'_1 := B_+(x', \xi').$$

Since  $\pi_\Gamma(x', \xi')$  is in the trapped set,  $x'$  lies in the convex core, which is the octagon used in the construction of  $\mathcal{F}_\ell$ . Thus there exists a constant  $C_2 > 0$  depending only on  $\mathcal{K}$  such that

$$C_2^{-1} \leq \mathcal{P}(x', y'_0) \leq C_2.$$

Applying (7.8), we see that

$$\begin{aligned} & (C_1 C_2)^{-\delta} r^\delta \mu_\delta(\mathcal{G}(y'_0, \Lambda_\ell) \cap B(\mathcal{G}(y'_0, y'_1), C_1/C_2)) \\ & \leq \mu_\delta(\mathcal{G}(y_0, \Lambda_\ell) \cap B(\mathcal{G}(y_0, y_1), r)) \\ & \leq (C_2/C_1)^\delta r^\delta \mu_\delta(\mathcal{G}(y'_0, \Lambda_\ell) \cap B(\mathcal{G}(y'_0, y'_1), C_1 C_2)). \end{aligned}$$

Therefore, in order to prove (7.15) it is enough to verify the inequalities

$$\begin{aligned} \mu_\delta(\mathcal{G}(y'_0, \Lambda_\ell) \cap B(\mathcal{G}(y'_0, y'_1), C_1 C_2)) & \leq C c_\ell, \\ \mu_\delta(\mathcal{G}(y'_0, \Lambda_\ell) \cap B(\mathcal{G}(y'_0, y'_1), C_1/C_2)) & \geq C^{-1} c_\ell \end{aligned}$$

for some constants  $C$  depending only on  $\mathcal{K}$  and  $c_\ell$  depending on  $\ell$ .

We have  $y'_0, y'_1 \in \Lambda_\ell$ , thus  $y'_0 \in \overline{D_j}, y'_1 \in \overline{D_k}$  for some  $j, k \in \{1, 2, 3, 4\}$ . Moreover,  $x' \in \mathcal{F}_\ell$  implies that  $j \neq k$ . Thus  $|y'_0 - y'_1|$  is bounded away from zero uniformly in  $\ell \in \mathcal{K}$ . Since  $\mathcal{G}$  is a smooth map away from the diagonal, we see that it is enough to prove the inequalities

$$\mu_\delta(\{y \in \Lambda_\ell \mid \mathcal{G}(y'_0, y) \in B(\mathcal{G}(y'_0, y'_1), C_1 C_2)\}) \leq C c_\ell, \quad (7.20)$$

$$\mu_\delta(\{y \in \Lambda_\ell \mid \mathcal{G}(y'_0, y) \in B(\mathcal{G}(y'_0, y'_1), C_1/C_2)\}) \geq C^{-1} c_\ell. \quad (7.21)$$

We put  $c_\ell := \mu_\delta(\Lambda_\ell)$ , then (7.20) follows automatically. To show (7.21), we note that for  $C_1$  large enough depending on  $\mathcal{K}$ , the set on the left-hand side contains  $\Lambda_\ell \cap \overline{D_k}$ . It remains to use Lemma 7.9.  $\square$

## APPENDIX A. CALCULATIONS ON HYPERBOLIC QUOTIENTS

In this appendix, we prove several technical lemmas from §4 and §7 regarding the geometry of convex co-compact hyperbolic quotients  $M = \Gamma \backslash \mathbb{H}^n$ . Here

$$\Gamma \subset G := \text{PSO}(1, n)$$

is a group of hyperbolic isometries.

A useful tool is the *coframe bundle*  $F^*M$  whose points have the form  $(x, \xi_1, \dots, \xi_n)$ , where  $x \in M$  and  $\xi_1, \dots, \xi_n \in T_x^*M$  form a positively oriented orthonormal basis. We identify  $F^*M$  with  $\Gamma \backslash G$  by the diffeomorphism

$$[\gamma] \in \Gamma \backslash G \mapsto \pi_\Gamma^F(\gamma(0), d\gamma(0)^{-T} \cdot dx_1, \dots, d\gamma(0)^{-T} \cdot dx_n)$$

where we use the ball model for  $\mathbb{H}^n$  and

$$\pi_\Gamma^F : F^*\mathbb{H}^n \rightarrow F^*M$$

is the covering map. Using this identification, we obtain a right action of  $G$  on  $F^*M$ . Then each element of the Lie algebra of  $G$ , viewed as a left invariant vector field, induces a vector field on  $F^*M$ . We use the induced vector fields

$$U_1^\pm, \dots, U_{n-1}^\pm, \quad \tilde{X}, \quad R_{ab}, \quad 2 \leq a, b \leq n \quad (\text{A.1})$$

on  $F^*M$  defined in [DFG, §3.2] ( $\tilde{X}$  is denoted by simply  $X$  in [DFG]). The following commutation relations hold [DFG, (3.8)]:

$$[\tilde{X}, U_j^\pm] = \pm U_j^\pm, \quad [\tilde{X}, R_{ab}] = 0, \quad 2 \leq a, b \leq n. \quad (\text{A.2})$$

The vector field  $\tilde{X}$  projects down to  $X$  by  $\pi_S$ ; in fact,  $\tilde{X}$  is the generator of parallel transport along geodesics. The vector fields  $R_{ab}$  generate rotations of the vectors  $(\xi_2, \dots, \xi_n)$ ; in fact, the kernel of  $d\pi_S$  is spanned by  $(R_{ab})$  where  $\pi_S$  is the following submersion:

$$\pi_S : F^*M \rightarrow S^*M, \quad (x, \xi_1, \dots, \xi_n) \mapsto (x, \xi_1). \quad (\text{A.3})$$

*Proof of Lemma 4.2.* We show the first statement, the second one is proved similarly. Since  $L_s$  is an integrable foliation, induction on  $m+k$  shows that the order of applying vector fields in (3.1) does not matter for establishing the bound. By (4.3), it suffices to prove the bounds

$$\begin{aligned} \sup_{T^*M \setminus 0} |\chi_2 Y_1 \dots Y_m Z_1 \dots Z_k (\xi \cdot \partial_\xi)^i X^j (\chi_1 \circ e^{tX})| &\leq Ch^{-\rho k}, \\ Y_1, \dots, Y_m \in C^\infty(T^*M \setminus 0; E_s), \quad Z_1, \dots, Z_k \in C^\infty(T^*M \setminus 0; E_u). \end{aligned} \quad (\text{A.4})$$

Since  $X$  commutes with itself and with  $\xi \cdot \partial_\xi$ , we have

$$(\xi \cdot \partial_\xi)^i X^j (\chi_1 \circ e^{tX}) = ((\xi \cdot \partial_\xi)^i X^j \chi_1) \circ e^{tX}$$

and thus we may assume that  $i = j = 0$ . Since  $e^{tX}$  is a homogeneous flow and  $E_s, E_u$  are homogeneous foliations, we may restrict to the cosphere bundle  $S^*M$ .

On  $F^*M$ , we have (see [DFG, §3.3])

$$\pi_S^* E_s = \text{span}(\{U_j^+\} \cup \{R_{ab}\}), \quad \pi_S^* E_u = \text{span}(\{U_j^-\} \cup \{R_{ab}\}),$$

(The fields  $U_j^\pm$  do not project down to vector fields on  $S^*M$  for  $n \geq 3$ , hence the use of the coframe bundle.)

Then (A.4) reduces to the following bound on  $F^*M$ :

$$\sup_{F^*M} |(\pi_S^* \chi_2) U_{j_1}^+ \dots U_{j_m}^+ U_{r_1}^- \dots U_{r_k}^- ((\pi_S^* \chi_1) \circ e^{t\tilde{X}})| \leq Ch^{-\rho k}.$$

Using (A.2), we see that

$$U_{j_1}^+ \dots U_{j_m}^+ U_{r_1}^- \dots U_{r_k}^- ((\pi_S^* \chi_1) \circ e^{t\tilde{X}}) = e^{(k-m)t} (U_{j_1}^+ \dots U_{j_m}^+ U_{r_1}^- \dots U_{r_k}^- (\pi_S^* \chi_1)) \circ e^{t\tilde{X}}$$

and the proof is finished as  $e^{(k-m)t} \leq e^{kt} \leq h^{-\rho k}$  when  $t \in [0, \rho \log(1/h)]$ .  $\square$

*Proof of (4.12).* We prove the second statement, the first one is proved similarly. Assume first that  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$  and  $\pi_\Gamma(x, \xi) \in \Gamma_-$ , where  $\pi_\Gamma$  is defined in (4.4) and descends to a covering map  $\mathbb{H}^n \rightarrow M$  (also called  $\pi_\Gamma$ ). Let  $x(t) \in \mathbb{H}^n$  be the projection of  $e^{tX}(x, \xi)$  to the base. Since  $(x, \xi) \in \Gamma_-$ , the curve  $\pi_\Gamma(x(t))$  stays in a compact subset of  $M$  when  $t \geq 0$ ; therefore, there exists a sequence  $t_j \rightarrow \infty$  and  $x_\infty \in \mathbb{H}^n$  such that

$$\pi_\Gamma(x(t_j)) \rightarrow \pi_\Gamma(x_\infty) \quad \text{as } j \rightarrow \infty.$$

Take  $\gamma_j \in \Gamma$  such that  $d_{\mathbb{H}^n}(x(t_j), \gamma_j.x_\infty) \rightarrow 0$  as  $j \rightarrow \infty$ . Since  $x(t_j) \rightarrow B_+(x, \xi)$  in the topology of the closed disk model  $\overline{\mathbb{H}^n}$ , we have  $\gamma_j.x_\infty \rightarrow B_+(x, \xi)$  in  $\overline{\mathbb{H}^n}$  and thus  $B_+(x, \xi) \in \Lambda_\Gamma$  by (4.11).

Assume now that  $\pi_\Gamma(x, \xi) \notin \Gamma_-$ . Then (see for instance [DyGu14a, §7.1 and Appendix A.1]) the trajectory  $\pi_\Gamma(x(t))$  converges in the compactified space  $\overline{M}$  to some point  $y \in \partial\overline{M}$  on the conformal boundary. The point  $B_+(x, \xi)$  projects to  $y$  by an extension of  $\pi_\Gamma$  to the conformal boundary and there exists a neighborhood  $U$  of  $B_+(x, \xi)$  in  $\overline{\mathbb{H}^n}$  such that  $\pi_\Gamma$  is a local diffeomorphism on  $U$ . This implies that  $B_+(x, \xi)$  cannot be in the limit set  $\Lambda_\Gamma$ , finishing the proof.  $\square$

Next, we prove Lemma 4.3. We will use the following

**Lemma A.1.** *Assume that  $(x, \xi_1, \dots, \xi_n) \in F^*\mathbb{H}^n$ ,  $y \in \mathbb{S}^{n-1}$ , and*

$$y \neq B_-(x, \xi_1). \tag{A.5}$$

*Then there exists  $\mathbf{s} = (s_1, \dots, s_{n-1}) \in \mathbb{R}^{n-1}$  such that, putting (see (A.1))*

$$e^{\mathbf{s}U^-} := \exp(s_1U_1^- + \dots + s_{n-1}U_{n-1}^-) : F^*\mathbb{H}^n \rightarrow F^*\mathbb{H}^n, \tag{A.6}$$

*we have (with  $\pi_S$  defined in (A.3))*

$$B_+(\pi_S(e^{\mathbf{s}U^-}(x, \xi_1, \dots, \xi_n))) = y. \tag{A.7}$$

*Moreover, if  $(x, \xi_1, y)$  varies in some fixed compact subset of  $S^*\mathbb{H}^n \times \mathbb{S}^{n-1}$  satisfying (A.5), then we may choose  $s_1, \dots, s_{n-1} \in [-C, C]$  where  $C$  is independent of  $x, \xi_1, \dots, \xi_n, y$ .*

*Proof.* We use the hyperboloid model of  $\mathbb{H}^n$ , see [DFG, §3.1]. In particular, we consider elements of  $G = \text{PSO}(1, n)$  as linear automorphisms of the Minkowski space  $\mathbb{R}^{1, n}$ . If  $e_0, \dots, e_n$  is the canonical basis of  $\mathbb{R}^{1, n}$  and  $\gamma \in G$  is defined by

$$\gamma.(e_0, \dots, e_n) = (x, \xi_1, \dots, \xi_n),$$

then a direct calculation using [DFG, (3.7) and (3.16)] shows that for some  $q > 0$ ,

$$(q, qB_+(\pi_S(e^{\mathbf{s}U^-}(x, \xi_1, \dots, \xi_n)))) = \gamma.(1 + |\mathbf{s}|^2, 1 - |\mathbf{s}|^2, -2s_1, \dots, -2s_{n-1})$$

where both sides of the equation are vectors in the positive half of the light cone in  $\mathbb{R}^{1,n}$ . The vector  $\gamma^{-1} \cdot (1, y)$  lies in the positive half of the light cone, therefore we have for some  $r > 0$  and  $v \in \mathbb{S}^{n-1}$ ,

$$\gamma^{-1} \cdot (1, y) = (r, rv).$$

By [DFG, (3.16)], we see that (A.5) implies that

$$v \neq (-1, 0, \dots, 0). \quad (\text{A.8})$$

To obtain (A.7), we have to find  $\mathbf{s}$  such that

$$v_1 = \frac{1 - |\mathbf{s}|^2}{1 + |\mathbf{s}|^2}; \quad v_{j+1} = -\frac{2s_j}{1 + |\mathbf{s}|^2}, \quad 1 \leq j \leq n-1.$$

The existence of such  $\mathbf{s}$  follows directly from (A.8), and the fact that  $\mathbf{s}$  can be chosen bounded uniformly in  $(x, \xi_1, \dots, \xi_n, y)$  follows immediately because under (A.5), the distance from  $v$  to  $(-1, 0, \dots, 0)$  is bounded away from zero.  $\square$

*Proof of Lemma 4.3.* Assume that  $t \geq 0$  and

$$(x, \xi) \in V, \quad \pi_\Gamma(e^{tX}(x, \xi)) \in \pi_\Gamma(V). \quad (\text{A.9})$$

(The case of propagation for negative time is handled similarly.) Since  $e^{tX}$  is a homogeneous flow, we may additionally assume that  $|\xi|_g = 1$ .

By (A.9), there exists  $\gamma' \in \Gamma$  such that  $\gamma' \cdot e^{tX}(x, \xi) \in V$ . Since  $K = \Gamma_+ \cap \Gamma_- \neq \emptyset$ , by (4.12)  $\Lambda_\Gamma$  has at least two elements. Then we may choose  $y \in \Lambda_\Gamma$  depending on  $\gamma'$  and  $(x, \xi)$  such that for some  $\varepsilon > 0$  independent of  $(x, \xi)$ ,

$$d(B_-(\gamma' \cdot (x, \xi)), y) > \varepsilon.$$

Choose  $\Xi := (\xi_2, \dots, \xi_n)$  such that  $(x, \xi, \Xi) \in F^*\mathbb{H}^n$ . Applying Lemma A.1 to  $\gamma' \cdot e^{t\tilde{X}}(x, \xi, \Xi)$ , we see that there exists a constant  $C$  independent of  $x, \xi, \Xi$  such that

$$B_+(\pi_S(e^{sU^-}(\gamma' \cdot e^{t\tilde{X}}(x, \xi, \Xi)))) = y \quad (\text{A.10})$$

for some  $\mathbf{s} \in [-C, C]^{n-1}$ . Rewriting (A.10) using (A.2) and since  $B_+$  stays constant along the geodesic flow, we get

$$B_+(\pi_S(\exp(e^{-t}\mathbf{s}U^-)(\gamma' \cdot (x, \xi, \Xi)))) = y.$$

It follows that

$$B_+(\pi_S(\exp(e^{-t}\mathbf{s}U^-)(x, \xi, \Xi))) = (\gamma')^{-1} \cdot y \in \Lambda_\Gamma \quad (\text{A.11})$$

where  $(\gamma')^{-1} \in \Gamma$  acts on the conformal infinity  $\mathbb{S}^{n-1}$  as in [DFG, §3.5] and  $\Lambda_\Gamma$  is invariant by this action (as can be seen from either (4.11) or (4.12)). It follows from (A.11) that  $d(B_+(x, \xi), \Lambda_\Gamma) \leq Ce^{-t}$  for some constant  $C$ , finishing the proof.  $\square$

*Proof of Lemma 4.7.* Consider the generating function

$$\mathcal{S} \in C^\infty(\mathbb{H}_x^n \times \mathbb{R}_w^+ \times \mathbb{S}_y^{n-1}; \mathbb{R}), \quad \mathcal{S}(x, w, y) = w \log \mathcal{P}(x, y). \quad (\text{A.12})$$

Let  $\xi, \theta, \eta$  be the momenta on  $T^*(\mathbb{H}^n \times \mathbb{R}^+ \times \mathbb{S}^{n-1})$  corresponding to  $x, w, y$ . We claim that the following two statements are equivalent for each choice of  $(x, w, y, \xi, \theta, \eta)$ :

$$\xi = \mp \partial_x \mathcal{S}, \quad \theta = \pm \partial_w \mathcal{S}, \quad \eta = \pm \partial_y \mathcal{S}; \quad (\text{A.13})$$

$$w = p(x, \xi), \quad y = B_\mp(x, \xi), \quad \theta = \pm \log \mathcal{P}(x, B_\mp(x, \xi)), \quad \eta = \pm G_\mp(x, \xi). \quad (\text{A.14})$$

We compute

$$\partial_x \mathcal{S}(x, w, y) = -2w \left( \frac{x}{1 - |x|^2} + \frac{x - y}{|x - y|^2} \right), \quad \partial_w \mathcal{S}(x, w, y) = \log \mathcal{P}(x, y),$$

and the tangent vector in  $T_y \mathbb{S}^{n-1} \subset \mathbb{R}^n$  identified with  $\partial_y \mathcal{S}(x, w, y)$  by the round metric on  $\mathbb{S}^{n-1}$  is

$$\zeta(x, w, y) = 2w \frac{x - (x \cdot y)y}{|x - y|^2}. \quad (\text{A.15})$$

For  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$ , put

$$\Phi_\pm(x, \xi) = p(x, \xi) \mathcal{P}(x, B_\pm(x, \xi)).$$

We calculate (using for example [DFG, §3.4]; note that in that paper,  $x$  denotes a point in the hyperboloid model and  $y$  a point in the ball model; we relate the two models by [DFG, (3.2)], identify tangent and cotangent vectors by the metric and extend the formulas from the cosphere bundle to the whole  $T^*\mathbb{H}^{n+1} \setminus 0$  by homogeneity)

$$\begin{aligned} p(x, \xi) &= \frac{1 - |x|^2}{2} |\xi|, \\ \Phi_\pm(x, \xi) &= \frac{1 + |x|^2}{2} |\xi| \pm x \cdot \xi, \end{aligned} \quad (\text{A.16})$$

$$\Phi_\pm(x, \xi) B_\pm(x, \xi) = (|\xi| \pm x \cdot \xi) x \pm \frac{1 - |x|^2}{2} \xi.$$

We now show that (A.13) and (A.14) are equivalent. Assume first that (A.13) holds. Then

$$|\xi| = \frac{2w}{1 - |x|^2}, \quad x \cdot \xi = \mp w \left( \frac{1 - |x|^2}{|x - y|^2} - \frac{1 + |x|^2}{1 - |x|^2} \right).$$

It follows that  $p(x, \xi) = w$ ,  $\Phi_\mp(x, \xi) = w \mathcal{P}(x, y)$ , and  $B_\mp(x, \xi) = y$ , which gives the first three equalities in (A.14). Next,

$$\Phi_\pm(x, \xi) = 2w \frac{1 + |x|^2}{1 - |x|^2} - w \mathcal{P}(x, y),$$

$$\Phi_\pm(x, \xi) B_\pm(x, \xi) = \frac{4w}{1 - |x|^2} x - w \mathcal{P}(x, y) y,$$

$$\Phi_\pm(x, \xi) (B_+(x, \xi) \cdot B_-(x, \xi)) = \Phi_\mp - \frac{2w}{\mathcal{P}(x, y)}.$$

From here we see that  $G_{\mp}(x, \xi) = \zeta(x, w, y)$ , with  $\zeta$  defined in (A.15); this finishes the proof of (A.14).

Assume now that (A.14) holds. We have

$$\begin{aligned}\Phi_{\mp}(x, \xi)(x \cdot y) &= \mp \frac{1 + |x|^2}{2}(x \cdot \xi) + |\xi| \cdot |x|^2, \\ \Phi_{\mp}(x, \xi)|x - y|^2 &= \frac{(1 - |x|^2)^2}{2}|\xi|, \\ \Phi_{\mp}(x, \xi)(x - y) &= -\frac{1 - |x|^2}{2}(|\xi|x \mp \xi), \\ \frac{2\Phi_{\mp}(x, \xi)^2}{1 - |x|^2}(x - (x \cdot y)y) &= \left( \frac{1 - |x|^2}{2}|\xi|^2 + (x \cdot \xi)^2 \mp |\xi|(x \cdot \xi) \right) x \\ &\quad - \left( \frac{1 + |x|^2}{2}(x \cdot \xi) \mp |\xi| \cdot |x|^2 \right) \xi.\end{aligned}$$

It follows that  $w\mathcal{P}(x, y) = \Phi_{\mp}(x, \xi)$ ,  $\partial_x \mathcal{S}(x, w, y) = \mp \xi$ , and  $\zeta(x, w, y) = G_{\mp}(x, \xi)$ ; this gives (A.13).

The equivalence of (A.13) and (A.14) implies that  $\varkappa^{\pm}$  is a local symplectomorphism, since it has a generating function  $\mp \mathcal{S}$ . Moreover,  $\varkappa^{\pm}$  is exact, as  $\mp \mathcal{S}(x, w, y)$  is an antiderivative, see (2.6). To show that  $\varkappa^{\pm}$  is a global symplectomorphism, it remains to prove that for each  $(w, y, \theta, \eta) \in T^*(\mathbb{R}^+ \times \mathbb{S}^{n-1})$ , there exists unique  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$  such that  $\varkappa^{\pm}(x, \xi) = (w, y, \theta, \eta)$ . This is a direct consequence of (4.44) and the geometry of the hyperbolic space: the values of  $w, \theta$  determine uniquely  $p(x, \xi)$ ,  $\Phi_{\mp}(x, \xi)$  and then  $y, \eta$  determine uniquely  $B_+(x, \xi)$ ,  $B_-(x, \xi)$ ; then  $B_+(x, \xi)$ ,  $B_-(x, \xi)$  determine the normalized geodesic passing through  $(x, \xi)$  and  $\Phi_{\mp}(x, \xi)$ ,  $p(x, \xi)$  determine the unique point  $(x, \xi)$  on that geodesic.  $\square$

*Proof of Lemma 4.8.* By (4.45), (4.47) is equivalent to the fact that

$$(w, y, \theta, \eta) = \varkappa^-(x, \xi), \quad (w, y', \theta', \eta') = \varkappa^+(x, \xi) \quad (\text{A.17})$$

for some  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$ , which by (4.44) is equivalent to

$$\begin{aligned}\theta' &= \theta + \log(1 + |\eta|^2/w^2), \\ y' &= \frac{(|\eta|^2 - w^2)y - 2w\eta}{|\eta|^2 + w^2}, \\ \eta' &= \frac{2w|\eta|^2 y + (|\eta|^2 - w^2)\eta}{|\eta|^2 + w^2}.\end{aligned} \quad (\text{A.18})$$

Here we identified vectors and covectors on  $\mathbb{S}^{n-1}$  by the round metric and used the following identities true for  $(x, \xi) \in T^*\mathbb{H}^n \setminus 0$  and  $(y, y') \in \mathbb{S}_\Delta^{n-1}$ :

$$\mathcal{P}(x, B_+(x, \xi))\mathcal{P}(x, B_-(x, \xi))(1 - B_+(x, \xi) \cdot B_-(x, \xi)) = 2, \quad (\text{A.19})$$

$$|\mathcal{G}(y, y')|^2 = \frac{1 + y \cdot y'}{1 - y \cdot y'}. \quad (\text{A.20})$$

We next see that (4.48) is equivalent to

$$\begin{aligned} \theta' &= \theta + \log \frac{4}{|y - y'|^2}, \\ \eta &= -\frac{2w(y' - (y \cdot y')y)}{|y - y'|^2}, \\ \eta' &= \frac{2w(y - (y \cdot y')y')}{|y - y'|^2}. \end{aligned} \quad (\text{A.21})$$

A direct calculation shows that (A.18) and (A.21) are equivalent; therefore, (4.47) and (4.48) are equivalent as well. By the proof of Lemma 4.7, the antiderivatives for  $\varkappa^+$  and  $(\varkappa^-)^{-1}$  are given by  $-w \log \mathcal{P}(x, y')$  and  $-w \log \mathcal{P}(x, y)$ , where  $x$  is defined by (A.17); their sum is then

$$-w \log \frac{2}{1 - y \cdot y'} = \Theta(w, y, y').$$

□

*Proof of Lemma 7.2.* We use the coframe bundle  $F^*\mathbb{H}^n$  introduced at the beginning of this appendix. Let  $(x, \xi) \in S^*\mathbb{H}^n$ ,  $\eta \in E_u(x, \xi)$ , and choose  $\Xi = (\xi_2, \dots, \xi_n)$  such that  $(x, \xi, \Xi) \in F^*\mathbb{H}^n$ .

As in the proof of Lemma A.1, we use the hyperboloid model of  $\mathbb{H}^n$  and take  $\gamma \in G$  which maps the standard frame to  $(x, \xi, \Xi)$ . Then a direct calculation using [DFG, (3.7)] shows that for each  $\mathbf{s} \in \mathbb{R}^{n-1}$ , we have

$$\pi_S(e^{\mathbf{s}U^-}(x, \xi, \Xi)) = \left( \gamma \cdot \left( 1 + \frac{|\mathbf{s}|^2}{2}, -\frac{|\mathbf{s}|^2}{2}, -\mathbf{s} \right), \gamma \cdot \left( \frac{|\mathbf{s}|^2}{2}, 1 - \frac{|\mathbf{s}|^2}{2}, -\mathbf{s} \right) \right) \quad (\text{A.22})$$

where  $e^{\mathbf{s}U^-} : F^*\mathbb{H}^n \rightarrow F^*\mathbb{H}^n$  is defined in (A.6) and we consider elements of  $S^*\mathbb{H}^n$  as pairs of vectors in the Minkowski space.

Take  $\mathbf{s} \in \mathbb{R}^{n-1}$  such that

$$\partial_r|_{r=0} \pi_S(e^{r\mathbf{s}U^-}(x, \xi, \Xi)) = \eta. \quad (\text{A.23})$$

As follows from a direct calculation using (A.22) and [DFG, (3.14)], such  $\mathbf{s}$  exists, is unique, and depends linearly on  $\eta$ ; moreover

$$|\mathbf{s}| = |\eta| \quad (\text{A.24})$$

with  $|\eta|$  defined in the paragraph preceding (4.5).

With  $\mathbf{s}$  solving (A.23), we put

$$\mathcal{U}_-(x, \xi, \eta) := \pi_S(e^{sU_-}(x, \xi, \Xi)). \quad (\text{A.25})$$

Then  $\mathcal{U}_-(x, \xi, \eta)$  does not depend on the choice of  $\Xi$ . Indeed, take another  $\Xi'$  such that  $(x, \xi, \Xi') \in F^*\mathbb{H}^n$ , and let  $\gamma' \in G$  map the standard basis of the Minkowski space to  $(x, \xi, \Xi')$ . Then  $\gamma' = \gamma\tilde{\gamma}$  where  $\tilde{\gamma}$  lies in the subgroup  $\text{SO}(n-1) \subset G = \text{PSO}(1, n)$  consisting of isometries of the Minkowski space which fix the first two basis vectors. Viewing  $\tilde{\gamma}$  as an isometry on  $\mathbb{R}^{n-1}$ , we get from (A.22)

$$\pi_S(e^{sU_-}(x, \xi, \Xi')) = \left( \gamma \cdot \left( 1 + \frac{|\mathbf{s}|^2}{2}, -\frac{|\mathbf{s}|^2}{2}, -\tilde{\gamma} \cdot \mathbf{s} \right), \gamma \cdot \left( \frac{|\mathbf{s}|^2}{2}, 1 - \frac{|\mathbf{s}|^2}{2}, -\tilde{\gamma} \cdot \mathbf{s} \right) \right).$$

Then  $s' := (\tilde{\gamma})^{-1} \cdot \mathbf{s}$  solves (A.23) with  $\Xi$  replaced by  $\Xi'$ , and the right-hand sides of (A.25) for  $\Xi, \mathbf{s}$  and for  $\Xi', \mathbf{s}'$  are equal.

The fact that  $\mathcal{U}_-$  commutes with the action of  $G$  follows immediately from its construction and the fact that  $U_i^-$  are  $G$ -invariant vector fields on  $F^*\mathbb{H}^n$ . The identity (7.2) follows directly from (A.23) and [DFG, (3.16)].

To see (7.3), note that by (1.19) and (A.19)

$$\mathcal{G}(B_-(x, \xi), B_+(x, \xi)) = \frac{\Phi_+(x, \xi)\Phi_-(x, \xi)}{2} B_+(x, \xi) + \left( 1 - \frac{\Phi_+(x, \xi)\Phi_-(x, \xi)}{2} \right) B_-(x, \xi)$$

where  $\Phi_{\pm}(x, \xi) := \mathcal{P}(x, B_{\pm}(x, \xi))$ . Put  $(\tilde{x}, \tilde{\xi}) := \mathcal{U}_-(x, \xi, \eta)$ . By (A.23) and [DFG, (3.16)], we have  $\Phi_-(\tilde{x}, \tilde{\xi}) = \Phi_-(x, \xi)$ . Then

$$\begin{aligned} & \mathcal{G}(B_-(x, \xi), B_+(\tilde{x}, \tilde{\xi})) - \mathcal{G}(B_-(x, \xi), B_+(x, \xi)) \\ &= \frac{\Phi_-(x, \xi)}{2} (\Phi_+(\tilde{x}, \tilde{\xi})B_+(\tilde{x}, \tilde{\xi}) - \Phi_+(x, \xi)B_+(x, \xi)) \\ & \quad - \frac{\Phi_+(\tilde{x}, \tilde{\xi}) - \Phi_+(x, \xi)}{2} \Phi_-(x, \xi)B_-(x, \xi). \end{aligned}$$

Now, by [DFG, (3.16)], we have in the hyperboloid model  $x \pm \xi = \Phi_{\pm}(x, \xi)(1, B_{\pm}(x, \xi))$ , therefore

$$\begin{aligned} & (0, \mathcal{G}(B_-(x, \xi), B_+(\tilde{x}, \tilde{\xi})) - \mathcal{G}(B_-(x, \xi), B_+(x, \xi))) \\ &= \frac{\Phi_-(x, \xi)}{2} (\tilde{x} + \tilde{\xi} - x - \xi) - \frac{\tilde{x}_0 + \tilde{\xi}_0 - x_0 - \xi_0}{2} \Phi_-(x, \xi)(1, B_-(x, \xi)) \\ &= \Phi_-(x, \xi)(\mathbf{w} - \mathbf{w}_0(1, B_-(x, \xi))), \quad \mathbf{w} = \gamma \cdot \left( \frac{|\mathbf{s}|^2}{2}, -\frac{|\mathbf{s}|^2}{2}, -\mathbf{s} \right) \in \mathbb{R}^{1, n} \end{aligned}$$

where the last identity follows from (A.22) and (A.25). Now, since  $\gamma \cdot (1, -1, 0) = \Phi_-(x, \xi)(1, B_-(x, \xi))$ , we have

$$\mathbf{w} = \frac{|\mathbf{s}|^2}{2} \Phi_-(x, \xi)(1, B_-(x, \xi)) - \gamma \cdot (0, 0, \mathbf{s})$$

therefore (7.3) holds with

$$(0, \mathcal{T}_-(x, \xi)\eta) = \mathbf{v} - \mathbf{v}_0(1, B_-(x, \xi)), \quad \mathbf{v} := -\gamma \cdot (0, 0, \mathbf{s}) \in \mathbb{R}^{1,n},$$

and the fact that  $\mathcal{T}_-$  is an isometry follows from (A.24) and the fact that  $(1, B_-(x, \xi))$  is a null vector in  $\mathbb{R}^{1,n}$  orthogonal to  $\mathbf{v}$ .  $\square$

**Acknowledgements.** We would like to thank Maciej Zworski for constant support and encouragement throughout the project; Dmitry Dolgopyat, Frédéric Naud, Stéphane Nonnenmacher, and Peter Sarnak for several insightful conversations on spectral gaps; Larry Guth, Alex Iosevich, Hans Parshall, and Tomasz Schoen for several helpful discussions regarding Theorem 6; Alexander Razborov and Jean Bourgain for comments and corrections to an earlier version of this manuscript; Colin Guillarmou and Peter Perry for a review of the history of hyperbolic scattering; Jeffrey Galkowski and Dmitry Jakobson for many interesting discussions regarding this project; and David Borthwick and Tobias Weich for providing the data for Figure 1(b).

During the period this research was conducted, the first author served as a Clay Research Fellow and the second author was supported by a NSF mathematical sciences postdoctoral research fellowship.

## REFERENCES

- [AnRo] James Anderson and André Rocha, *Analyticity of Hausdorff dimension of limit sets of Kleinian groups*, Ann. Acad. Sci. Fen. Math. **22**(1997), 349–364.
- [BWPSKZ] Sonja Barkhofen, Tobias Weich, Alexander Potzweit, Hans-Jürgen Stöckmann, Ulrich Kuhl, and Maciej Zworski, *Experimental observation of the spectral gap in microwave  $n$ -disk systems*, Phys. Rev. Lett. **110**(2013), 164102.
- [BLZ] Matthew Bond, Izabella Laba, and Joshua Zahl, *Quantitative visibility estimates for unrectifiable sets in the plane*, preprint, [arXiv:1306.5469](https://arxiv.org/abs/1306.5469).
- [BoMi] Jean-François Bony and Laurent Michel, *Microlocalization of resonant states and estimates of the residue of the scattering amplitude*, Comm. Math. Phys. **246**(2004), 375–402.
- [Bo07] David Borthwick, *Spectral theory of infinite-area hyperbolic surfaces*, Birkhäuser, 2007.
- [Bo14] David Borthwick, *Distribution of resonances for hyperbolic surfaces*, Experimental Math. **23**(2014), 25–45.
- [BoWe] David Borthwick and Tobias Weich, *Symmetry reduction of holomorphic iterated function schemes and factorization of Selberg zeta functions*, J. Spectr. Th. **6**(2016), 267–329.
- [BGS] Jean Bourgain, Alex Gamburd, and Peter Sarnak, *Generalization of Selberg’s 3/16 theorem and affine sieve*, Acta Math. **207**(2011), 255–290.
- [BuOl97] Ulrich Bunke and Martin Olbrich, *Fuchsian groups of the second kind and representations carried by the limit set*, Invent. Math. **127**(1997), 127–154.
- [BuOl99] Ulrich Bunke and Martin Olbrich, *Group cohomology and the singularities of the Selberg zeta function associated to a Kleinian group*, Ann. of Math. (2) **149**(1999), 627–689.
- [Co] Michel Coornaert, *Mesures de Patterson–Sullivan sur le bord d’un espace hyperbolique au sens de Gromov*, Pacific J. Math. **159**(1993), 241–270.

- [Da] Kiril Datchev, *Resonance free regions for nontrapping manifolds with cusps*, to appear in Analysis&PDE; [arXiv:1210.7736](#).
- [DaDy] Kiril Datchev and Semyon Dyatlov, *Fractal Weyl laws for asymptotically hyperbolic manifolds*, Geom. Funct. Anal. **23**(2013), 1145–1206.
- [DaSe] Guy David and Stephen Semmes, *Fractured fractals and broken dreams: Self-similar geometry through metric and measure*, Oxford University Press, 1997.
- [De] Jean-Marc Delort, *FBI transformation, second microlocalization, and semilinear caustics*, Springer, 1992.
- [Do] Dmitry Dolgopyat, *On decay of correlations in Anosov flows*, Ann. Math. (2) **147**(1998), 357–390.
- [Dy] Semyon Dyatlov, *Resonance projectors and asymptotics for  $r$ -normally hyperbolic trapped sets*, J. Amer. Math. Soc. **28**(2015), 311–381.
- [DFG] Semyon Dyatlov, Frédéric Faure, and Colin Guillarmou, *Power spectrum of the geodesic flow on hyperbolic manifolds*, Analysis&PDE **8**(2015), 923–1000.
- [DyGu14a] Semyon Dyatlov and Colin Guillarmou, *Microlocal limits of plane waves and Eisenstein functions*, Ann. de l'ENS (4) **47**(2014), 371–448.
- [DyGu14b] Semyon Dyatlov and Colin Guillarmou, *Pollicott–Ruelle resonances for open systems*, Ann. Henri Poincaré, published online; [arXiv:1410.5516](#).
- [DyZw16] Semyon Dyatlov and Maciej Zworski, *Dynamical zeta functions for Anosov flows via microlocal analysis*, Ann. Sc. Ec. Norm. Sup. **49**(2016), 543–577.
- [DyZw] Semyon Dyatlov and Maciej Zworski, *Mathematical theory of scattering resonances*, book in progress, <http://math.mit.edu/~dyatlov/res/>
- [Fr] Grigory Freĭman, *Foundations of a structural theory of set addition*, Translations of mathematical monographs **37**, AMS, 1973.
- [GaRi] Pierre Gaspard and Stuart Rice, *Scattering from a classically chaotic repeller*, J. Chem. Phys. **90**(1989), 2225–2241.
- [GrSj] Alain Grigis and Johannes Sjöstrand, *Microlocal analysis for differential operators: an introduction*, Cambridge University Press, 1994.
- [Gu] Colin Guillarmou, *Meromorphic properties of the resolvent on asymptotically hyperbolic manifolds*, Duke Math. J. **129**(2005), 1–37.
- [GuSt77] Victor Guillemin and Shlomo Sternberg, *Geometric asymptotics*, Mathematical Surveys **14**, AMS, 1977.
- [GuSt13] Victor Guillemin and Shlomo Sternberg, *Semi-classical analysis*, International Press, 2013.
- [GuZw] Laurent Guillopé and Maciej Zworski, *Polynomial bounds on the number of resonances for some complete spaces of constant negative curvature near infinity*, Asymp. Anal. **11**(1995), 1–22.
- [HöI] Lars Hörmander, *The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis*, Springer, 1990.
- [HöIII] Lars Hörmander, *The Analysis of Linear Partial Differential Operators III. Pseudo-Differential Operators*, Springer, 1994.
- [Ik] Mitsuru Ikawa, *Decay of solutions of the wave equation in the exterior of several convex bodies*, Ann. Inst. Fourier **38**(1988), 113–146.
- [JaNa] Dmitry Jakobson and Frédéric Naud, *On the critical line of convex co-compact hyperbolic surfaces*, Geom. Funct. Anal. **22**(2012), 352–368.
- [MOW] Michael Magee, Hee Oh, and Dale Winter, *Expanding maps and continued fractions*, preprint, [arXiv:1412.4284](#).

- [MaMe] Rafe Mazzeo and Richard Melrose, *Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature*, J. Funct. Anal. **75**(1987), 260–310.
- [Mc] Curtis McMullen, *Hausdorff dimension and conformal dynamics III: computation of dimension*, Amer. J. Math. **120**(1998), 691–721.
- [Na] Frédéric Naud, *Expanding maps on Cantor sets and analytic continuation of zeta functions*, Ann. de l'ENS (4) **38**(2005), 116–153.
- [NoZw09] Stéphane Nonnenmacher and Maciej Zworski, *Quantum decay rates in chaotic scattering*, Acta Math. **203**(2009), 149–233.
- [NoZw15] Stéphane Nonnenmacher and Maciej Zworski, *Decay of correlations for normally hyperbolic trapping*, Invent. Math. **200**(2015), 345–438.
- [OhWi] Hee Oh and Dale Winter, *Uniform exponential mixing and resonance free regions for convex cocompact congruence subgroups of  $SL_2(\mathbb{Z})$* , to appear in J. Amer. Math. Soc., [arXiv:1410.4401](https://arxiv.org/abs/1410.4401).
- [Pa75] Samuel James Patterson, *The Laplacian operator on a Riemann surface. I*, Compos. Math. **31**(1975), 83–107.
- [Pa76a] Samuel James Patterson, *The Laplacian operator on a Riemann surface. II*, Compos. Math. **32**(1976), 71–112.
- [Pa76b] Samuel James Patterson, *The limit set of a Fuchsian group*, Acta Math. **136**(1976), 241–273.
- [PaPe] Samuel James Patterson and Peter Perry, *The divisor of Selberg's zeta function for Kleinian groups*, Duke Math. J. **106**(2001), 321–390.
- [Pe87] Peter Perry, *The Laplace operator on a hyperbolic manifold I. Spectral and scattering theory*, J. Funct. Anal. **75**(1987), 161–187.
- [Pe89] Peter Perry, *The Laplace operator on a hyperbolic manifold. II. Eisenstein series and the scattering matrix*, J. Reine Angew. Math. **398**(1989), 67–91.
- [PeSt] Vesselin Petkov and Luchezar Stoyanov, *Analytic continuation of the resolvent of the Laplacian and the dynamical zeta function*, Anal. PDE **3**(2010), 427–489.
- [Pe] Giorgis Petridis, *New proofs of Plünnecke-type estimates for product sets in groups*, Combinatorica **32** (2012), 721–733.
- [Ra] John Ratcliffe, *Foundations of hyperbolic manifolds*, second edition, Springer, 2006.
- [Ru] Imre Ruzsa, *Sums of finite sets*, in *Number Theory: New York Seminar, 1996*, 281–293.
- [Sa] Tom Sanders, *The structure theory of set addition revisited*, Bull. Amer. Math. Soc. **50**(2013), 93–127.
- [SjZw] Johannes Sjöstrand and Maciej Zworski, *Fractal upper bounds on the density of semiclassical resonances*, Duke Math. J. **137**(2007), 381–459.
- [St11] Luchezar Stoyanov, *Spectra of Ruelle transfer operators for axiom A flows*, Nonlinearity **24**(2011), 1089–1120.
- [St12] Luchezar Stoyanov, *Non-integrability of open billiard flows and Dolgopyat-type estimates*, Erg. Theory Dyn. Syst. **32**(2012), 295–313.
- [Su] Dennis Sullivan, *The density at infinity of a discrete group of hyperbolic motions*, Publ. Math. de l'IHES **50**(1979), 171–202.
- [TaVu06] Terence Tao and Van Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics **105**, Cambridge University Press, 2006.
- [TaVu08] Terence Tao and Van Vu, *John-type theorems for generalized arithmetic progressions and iterated sumsets*, Adv. Math. **219**(2008), 428–449.
- [Va13a] András Vasy, *Microlocal analysis of asymptotically hyperbolic and Kerr–de Sitter spaces*, with an appendix by Semyon Dyatlov, Invent. Math. **194**(2013), 381–513.

- [Va13b] András Vasy, *Microlocal analysis of asymptotically hyperbolic spaces and high energy resolvent estimates*, Inverse Problems and Applications. Inside Out II, Gunther Uhlmann (ed.), MSRI publications **60**, Cambridge Univ. Press, 2013.
- [Zw99] Maciej Zworski, *Dimension of the limit set and the density of resonances for convex co-compact hyperbolic surfaces*, Invent. Math. **136**(1999), 353–409.
- [Zw] Maciej Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, AMS, 2012.
- [Zw15] Maciej Zworski, *Resonances for asymptotically hyperbolic manifolds: Vasy’s method revisited*, preprint, [arXiv:1511.03352](https://arxiv.org/abs/1511.03352).

*E-mail address:* dyatlov@math.mit.edu

*E-mail address:* jzahl@mit.edu

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVE, CAMBRIDGE, MA 02139