Analysis & PDE

Volume 11 No. 1 2018

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2018 is US $275/year for the electronic version, and $480/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.1

ANALYTIC TORSION, DYNAMICAL ZETA FUNCTIONS, AND THE FRIED CONJECTURE

SHU SHEN

We prove the equality of the analytic torsion and the value at zero of a Ruelle dynamical zeta function associated with an acyclic unitarily flat vector bundle on a closed locally symmetric reductive manifold. This solves a conjecture of Fried. This article should be read in conjunction with an earlier paper by Moscovici and Stanton.

1. Introduction

The purpose of this article is to prove the equality of the analytic torsion and the value at zero of a Ruelle dynamical zeta function associated with an acyclic unitarily flat vector bundle on a closed locally symmetric reductive manifold, which completes a gap in the proof given by Moscovici and Stanton [1991] and solves a conjecture of Fried [1987].

Let Z be a smooth closed manifold. Let F be a complex vector bundle equipped with a flat Hermitian metric g^F. Let $H^\bullet(Z,F)$ be the cohomology of sheaf of locally flat sections of F. We assume $H^\bullet(Z,F)=0$.

The Reidemeister torsion, introduced in [Reidemeister 1935], is a positive real number one obtains via the combinatorial complex with values in F associated with a triangulation of Z, which can be shown not to depend on the triangulation.

Let g^{TZ} be a Riemannian metric on TZ. Ray and Singer [1971] constructed the analytic torsion $T(F)$ as a spectral invariant of the Hodge Laplacian associated with g^{TZ} and g^F. They showed that if Z is an

MSC2010: 58J20, 58J52, 11F72, 11M36, 37C30.
Keywords: index theory and related fixed point theorems, analytic torsion, Selberg trace formula, dynamical zeta functions.

even-dimensional oriented manifold, then $T(F) = 1$. Moreover, if $\dim Z$ is odd, then $T(F)$ does not depend on the metric data.

Ray and Singer [1971] conjectured an equality between the Reidemeister torsion and the analytic torsion, which was later proved by Cheeger [1979] and Müller [1978]. Using the Witten deformation, Bismut and Zhang [1992] gave an extension of the Cheeger–Müller theorem which is valid for arbitrary flat vector bundles.

From the dynamical side, Milnor [1968b, Section 3] pointed out a remarkable similarity between the Reidemeister torsion and the Weil zeta function. A quantitative description of their relation was formulated by Fried [1986] when Z is a closed oriented hyperbolic manifold. Namely, he showed that the value at zero of the Ruelle dynamical zeta function, constructed using the closed geodesics in Z and the holonomy of F, is equal to $T(F)^2$. Fried [1987, p. 66, Conjecture] suggested that a similar result holds true for general closed locally homogeneous manifolds.

In this article, we prove the Fried conjecture for odd-dimensional[1] closed locally symmetric reductive manifolds. More precisely, we show that the dynamical zeta function is meromorphic on $\mathbb{C}$, holomorphic at 0, and that its value at 0 is equal to $T(F)^2$.

The proof of the above result by Moscovici and Stanton [1991], based on the Selberg trace formula and harmonic analysis on reductive groups, does not seem to be complete. We give the proper argument to make it correct. Our proof is based on the explicit formula given by Bismut [2011, Theorem 6.1.1] for semisimple orbital integrals.

The results contained in this article were announced in [Shen 2016]. See also Ma's talk [2017] at Séminaire Bourbaki for an introduction.

Now, we will describe our results in more detail, and explain the techniques used in their proofs.

1A. ***The analytic torsion.*** Let Z be a smooth closed manifold, and let F be a complex flat vector bundle on Z.

Let g^{TZ} be a Riemannian metric on TZ, and let g^F be a Hermitian metric on F. To g^{TZ} and g^F, we can associate an L^2-metric on $\Omega^\bullet(Z, F)$, the space of differential forms with values in F. Let $\square^Z$ be the Hodge Laplacian acting on $\Omega^\bullet(Z, F)$. By Hodge theory, we have a canonical isomorphism

$$\ker \square^Z \simeq H^\bullet(Z, F). \tag{1-1}$$

Let $(\square^Z)^{-1}$ be the inverse of $\square^Z$ acting on the orthogonal space to $\ker \square^Z$. Let $N^{\Lambda^\bullet(T^*Z)}$ be the number operator of $\Lambda^\bullet(T^*Z)$, i.e., multiplication by i on $\Omega^i(Z, F)$. Let $\mathrm{Tr_s}$ denote the supertrace. For $s \in \mathbb{C},\ \mathrm{Re}(s) > \frac{1}{2}\dim Z$, set

$$\theta(s) = -\mathrm{Tr_s}[N^{\Lambda^\bullet(T^*Z)}(\square^Z)^{-s}]. \tag{1-2}$$

By [Seeley 1967], $\theta(s)$ has a meromorphic extension to $\mathbb{C}$, which is holomorphic at $s = 0$. The analytic torsion is a positive real number given by

$$T(F) = \exp(\theta'(0)/2). \tag{1-3}$$

[1] The even-dimensional case is trivial.

Equivalently, $T(F)$ is given by the following weighted product of the zeta regularized determinants:

$$T(F) = \prod_{i=1}^{\dim Z} \det(\square^Z|_{\Omega^i(Z,F)})^{(-1)^i i/2}. \tag{1-4}$$

1B. ***The dynamical zeta function.*** Let us recall the general definition of the formal dynamical zeta function associated to a geodesic flow given in [Fried 1987, Section 5].

Let (Z, g^{TZ}) be a connected manifold with nonpositive sectional curvature. Let $\Gamma = \pi_1(Z)$ be the fundamental group of Z, and let $[\Gamma]$ be the set of the conjugacy classes of Γ. We identify $[\Gamma]$ with the free homotopy space of Z. For $[\gamma] \in [\Gamma]$, let $B_{[\gamma]}$ be the set of closed geodesics, parametrized by $[0, 1]$, in the class $[\gamma]$. The map $x_\bullet \in B_{[\gamma]} \to (x_0, \dot{x}_0/|\dot{x}_0|)$ induces an identification between $\coprod_{[\gamma]\in[\Gamma]-\{1\}} B_{[\gamma]}$ and the fixed points of the geodesic flow at time $t = 1$ acting on the unit tangent bundle SZ. Then, $B_{[\gamma]}$ is equipped with the induced topology, and is connected and compact. Moreover, all the elements in $B_{[\gamma]}$ have the same length $l_{[\gamma]}$. Also, the Fuller index $\mathrm{ind}_F(B_{[\gamma]}) \in \mathbb{Q}$ is well defined [Fried 1987, Section 4]. Given a finite-dimensional representation ρ of Γ, for $\sigma \in \mathbb{C}$, the formal dynamical zeta function is then defined by

$$R_\rho(\sigma) = \exp\biggl(\sum_{[\gamma]\in[\Gamma]-\{1\}} \mathrm{Tr}[\rho(\gamma)]\,\mathrm{ind}_F(B_{[\gamma]})e^{-\sigma l_{[\gamma]}}\biggr). \tag{1-5}$$

Note that our definition is the inverse of the one introduced by Fried [1987, p. 51].

The Fuller index can be made explicit in many case. If $[\gamma] \in [\Gamma]-\{1\}$, the group $\mathbb{S}^1$ acts locally freely on $B_{[\gamma]}$ by rotation. Assume that the $B_{[\gamma]}$ are smooth manifolds. This is the case if (Z, g^{TZ}) has a negative sectional curvature or if Z is locally symmetric. Then $\mathbb{S}^1\backslash B_{[\gamma]}$ is an orbifold. Let $\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]}) \in \mathbb{Q}$ be the orbifold Euler characteristic [Satake 1957]. Denote by

$$m_{[\gamma]} = \bigl|\ker(\mathbb{S}^1 \to \mathrm{Diff}(B_{[\gamma]}))\bigr| \in \mathbb{N}^* \tag{1-6}$$

the multiplicity of a generic element in $B_{[\gamma]}$. By [Fried 1987, Lemma 5.3], we have

$$\mathrm{ind}_F(B_{[\gamma]}) = \frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}. \tag{1-7}$$

By (1-5) and (1-7), the formal dynamical zeta function is then given by

$$R_\rho(\sigma) = \exp\biggl(\sum_{[\gamma]\in[\Gamma]-\{1\}} \mathrm{Tr}[\rho(\gamma)]\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}e^{-\sigma l_{[\gamma]}}\biggr). \tag{1-8}$$

We will say that the formal dynamical zeta function is well defined if $R_\rho(\sigma)$ is holomorphic for $\mathrm{Re}(\sigma) \gg 1$ and extends meromorphically to $\sigma \in \mathbb{C}$.

Observe that if (Z, g^{TZ}) is of negative sectional curvature, then $B_{[\gamma]} \simeq \mathbb{S}^1$ and

$$\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]}) = 1. \tag{1-9}$$

In this case, $R_\rho(\sigma)$ was recently shown to be well defined by Giulietti, Liverani and Pollicott [Giulietti et al. 2013] and Dyatlov and Zworski [2016]. Moreover, Dyatlov and Zworski [2017] showed that if (Z, g^{TZ}) is a negatively curved surface, the order of the zero of $R_\rho(\sigma)$ at $\sigma = 0$ is related to the genus of Z.

1C. ***The Fried conjecture.*** Let us briefly recall the results in [Fried 1986]. Assume Z is an odd-dimensional connected orientable closed hyperbolic manifold. Take $r \in \mathbb{N}$. Let $\rho : \Gamma \to U(r)$ be a unitary representation of the fundamental group Γ. Let F be the unitarily flat vector bundle on Z associated to ρ.

Using the Selberg trace formula, Fried [1986, Theorem 3] showed that there exist explicit constants $C_\rho \in \mathbb{R}^*$ and $r_\rho \in \mathbb{Z}$ such that as $\sigma \to 0$,

$$R_\rho(\sigma) = C_\rho T(F)^2 \sigma^{r_\rho} + \mathcal{O}(\sigma^{r_\rho+1}). \tag{1-10}$$

Moreover, if $H^\bullet(Z, F) = 0$, then

$$C_\rho = 1, \quad r_\rho = 0, \tag{1-11}$$

so that

$$R_\rho(0) = T(F)^2. \tag{1-12}$$

Fried [1987, p. 66, Conjecture] suggested that the same holds true when Z is a general closed locally homogeneous manifold.

1D. ***The V-invariant.*** In this and in the following subsections, we give a formal proof of (1-12) using the V-invariant of Bismut and Goette [2004].

Let S be a closed manifold equipped with an action of a compact Lie group L, with Lie algebra $\mathfrak{l}$. If $a \in \mathfrak{l}$, let a^S be the corresponding vector field on S. Bismut and Goette [2004] introduced the V-invariant $V_a(S) \in \mathbb{R}$.

Let f be an a^S-invariant Morse–Bott function on S. Let $B_f \subset S$ be the critical submanifold. Since $a^S|_{B_f} \in TB_f$, $V_a(B_f)$ is also well defined. By [Bismut and Goette 2004, Theorem 4.10], $V_a(S)$ and $V_a(B_f)$ are related by a simple formula.

1E. ***Analytic torsion and the V-invariant.*** Let us argue formally. Let LZ be the free loop space of Z equipped with the canonical $\mathbb{S}^1$-action. Write $LZ = \coprod_{[\gamma]\in[\Gamma]} (LZ)_{[\gamma]}$ as a disjoint union of its connected components. Let a be the generator of the Lie algebra of $\mathbb{S}^1$ such that $\exp(a) = 1$. As explained in [Bismut 2005, Equation (0.3)], if F is a unitarily flat vector bundle on Z such that $H^\bullet(Z, F) = 0$, at least formally, we have

$$\log T(F) = -\sum_{[\gamma]\in[\Gamma]} \operatorname{Tr}[\rho(\gamma)] V_a((LZ)_{[\gamma]}). \tag{1-13}$$

Suppose that (Z, g^{TZ}) is an odd-dimensional connected closed manifold of nonpositive sectional curvature, and suppose that the energy functional

$$E : x_\bullet \in LZ \to \frac{1}{2}\int_0^1 |\dot{x}_s|^2\, ds \tag{1-14}$$

on LZ is Morse–Bott. The critical set of E is just $\coprod_{[\gamma]\in[\Gamma]} B_{[\gamma]}$, and all the critical points are local minima. Applying [Bismut and Goette 2004, Theorem 4.10] to the infinite-dimensional manifold $(LZ)_{[\gamma]}$ equipped with the $\mathbb{S}^1$-invariant Morse–Bott functional E, we have the formal identity

$$V_a((LZ)_{[\gamma]}) = V_a(B_{[\gamma]}). \tag{1-15}$$

Since $B_{[1]} \simeq Z$ is formed of the trivial closed geodesics, by the definition of the V-invariant,

$$V_a(B_{[1]}) = 0. \tag{1-16}$$

By [Bismut and Goette 2004, Proposition 4.26], if $[\gamma] \in [\Gamma] - \{1\}$, then

$$V_a(B_{[\gamma]}) = -\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1 \backslash B_{[\gamma]})}{2m_{[\gamma]}}. \tag{1-17}$$

By (1-13), (1-15)–(1-17), we get a formal identity

$$\log T(F) = \frac{1}{2} \sum_{[\gamma] \in [\Gamma] - \{1\}} \operatorname{Tr}[\rho(\gamma)] \frac{\chi_{\mathrm{orb}}(\mathbb{S}^1 \backslash B_{[\gamma]})}{m_{[\gamma]}}, \tag{1-18}$$

which is formally equivalent to (1-12).

1F. *The main result of the article.* Let G be a linear connected real reductive group [Knapp 1986, p. 3], and let θ be the Cartan involution. Let K be the maximal compact subgroup of G of the points of G that are fixed by θ. Let $\mathfrak{k}$ and $\mathfrak{g}$ be the Lie algebras of K and G, and let $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ be the Cartan decomposition. Let B be a nondegenerate bilinear symmetric form on $\mathfrak{g}$ which is invariant under the adjoint action of G and θ. Assume that B is positive on $\mathfrak{p}$ and negative on $\mathfrak{k}$. Set $X = G/K$. Then B induces a Riemannian metric g^{TX} on the tangent bundle $TX = G \times_K \mathfrak{p}$ such that X is of nonpositive sectional curvature.

Let $\Gamma \subset G$ be a discrete torsion-free cocompact subgroup of G. Set $Z = \Gamma \backslash X$. Then Z is a closed locally symmetric manifold with $\pi_1(Z) = \Gamma$. Recall that $\rho : \Gamma \to \mathrm{U}(r)$ is a unitary representation of Γ, and that F is the unitarily flat vector bundle on Z associated with ρ. The main result of this article gives the solution of the Fried conjecture for Z. In particular, this conjecture is valid for all the closed locally symmetric spaces of noncompact type.

Theorem 1.1. *Assume* $\dim Z$ *is odd. The dynamical zeta function* $R_\rho(\sigma)$ *is holomorphic for* $\operatorname{Re}(\sigma) \gg 1$ *and extends meromorphically to* $\sigma \in \mathbb{C}$. *Moreover*, *there exist explicit constants* $C_\rho \in \mathbb{R}^*$ *and* $r_\rho \in \mathbb{Z}$, *see* (7-75), *such that*, *when* $\sigma \to 0$,

$$R_\rho(\sigma) = C_\rho T(F)^2 \sigma^{r_\rho} + \mathcal{O}(\sigma^{r_\rho + 1}). \tag{1-19}$$

If $H^\bullet(Z, F) = 0$, *then*

$$C_\rho = 1, \quad r_\rho = 0, \tag{1-20}$$

so that

$$R_\rho(0) = T(F)^2. \tag{1-21}$$

Let $\delta(G)$ be the nonnegative integer defined by the difference between the complex ranks of G and K. Since $\dim Z$ is odd, $\delta(G)$ is odd. For $\delta(G) \neq 1$, Theorem 1.1 is originally due to Moscovici and Stanton [1991] and was recovered by Bismut [2011]. Indeed, it was proved in [Moscovici and Stanton 1991, Corollary 2.2, Remark 3.7] or [Bismut 2011, Theorem 7.9.3] that $T(F) = 1$ and $\chi_{\mathrm{orb}}(\mathbb{S}^1 \backslash B_{[\gamma]}) = 0$ for all $[\gamma] \in [\Gamma] - \{1\}$.

Remark that both of the above two proofs use the Selberg trace formula. However, in the evaluation of the geometric side of the Selberg trace formula and of orbital integrals, Moscovici and Stanton relied on Harish-Chandra's Plancherel theory, while Bismut used his explicit formula [2011, Theorem 6.1.1] obtained via the hypoelliptic Laplacian.

Our proof of Theorem 1.1 relies on Bismut's formula.

1G. *Our results on $R_\rho(\sigma)$.* Assume that $\delta(G)=1$. To show that $R_\rho(\sigma)$ extends as a meromorphic function on $\mathbb{C}$ when Z is hyperbolic, Fried [1986] showed that $R_\rho(\sigma)$ is an alternating product of certain Selberg zeta functions. Moscovici and Stanton's idea was to introduce the more general Selberg zeta functions and to get a similar formula for $R_\rho(\sigma)$.

Let us recall some facts about reductive group G with $\delta(G)=1$. In this case, there exists a unique (up to conjugation) standard parabolic subgroup $Q\subset G$ with Langlands decomposition $Q=M_QA_QN_Q$ such that $\dim A_Q=1$. Let $\mathfrak{m},\mathfrak{b},\mathfrak{n}$ be the Lie algebras of M_Q,A_Q,N_Q. Let $\alpha\in\mathfrak{b}^*$ be such that, for $a\in\mathfrak{b}$, $\mathrm{ad}(a)$ acts on $\mathfrak{n}$ as a scalar $\langle\alpha,a\rangle\in\mathbb{R}$ (see Proposition 6.3). Let M be the connected component of identity of M_Q. Then M is a connected reductive group with maximal compact subgroup $K_M=M\cap K$ and with Cartan decomposition $\mathfrak{m}=\mathfrak{p}_\mathfrak{m}\oplus\mathfrak{k}_\mathfrak{m}$. We have the identity of real K_M-representations

$$\mathfrak{p}\simeq\mathfrak{p}_\mathfrak{m}\oplus\mathfrak{b}\oplus\mathfrak{n}. \tag{1-22}$$

An observation due to Moscovici and Stanton is that $\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})\neq 0$ only if γ can be conjugated by an element of G into A_QK_M. For $\sigma\in\mathbb{C}$, we define the formal Selberg zeta function by

$$Z_j(\sigma)=\exp\Biggl(-\sum_{[\gamma]\in[\Gamma]-\{1\}}\mathrm{Tr}[\rho(\gamma)]\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}\frac{\mathrm{Tr}^{\Lambda^j(\mathfrak{n}^*)}[\mathrm{Ad}(k^{-1})]}{\bigl|\det(1-\mathrm{Ad}(e^ak^{-1}))|_{\mathfrak{n}\oplus\theta\mathfrak{n}}\bigr|^{\frac{1}{2}}}e^{-\sigma l_{[\gamma]}}\Biggr), \tag{1-23}$$

where $a\in\mathfrak{b}$, $k\in K_M$ are such that γ can be conjugated to e^ak^{-1}. We remark that $l_{[\gamma]}=|a|$. To show the meromorphicity of $Z_j(\sigma)$, Moscovici and Stanton tried to identify $Z_j(\sigma)$ with the geometric side of the zeta regularized determinant of the resolvent of some elliptic operator acting on some vector bundle on Z. However, the vector bundle used in [Moscovici and Stanton 1991], whose construction involves the adjoint representation of K_M on $\Lambda^i(\mathfrak{p}_\mathfrak{m}^*)\otimes\Lambda^i(\mathfrak{n}^*)$, does not live on Z, but only on $\Gamma\backslash G/K_M$.

We complete this gap by showing that such an object exists as a virtual vector bundle on Z in the sense of K-theory. More precisely, let $\mathrm{RO}(K),\mathrm{RO}(K_M)$ be the real representation rings of K and K_M. We can verify that the restriction $\mathrm{RO}(K)\to\mathrm{RO}(K_M)$ is injective. Note that $\mathfrak{p}_\mathfrak{m},\mathfrak{n}\in\mathrm{RO}(K_M)$. In Section 6C, using the classification theory of real simple Lie algebras, we show $\mathfrak{p}_\mathfrak{m},\mathfrak{n}$ are in the image of $\mathrm{RO}(K)$. For $0\leqslant j\leqslant\dim\mathfrak{n}$, let $E_j=E_j^+-E_j^-\in\mathrm{RO}(K)$ such that the following identity in $\mathrm{RO}(K_M)$ holds:

$$\Biggl(\sum_{i=0}^{\dim\mathfrak{p}_\mathfrak{m}}(-1)^i\Lambda^i(\mathfrak{p}_\mathfrak{m}^*)\Biggr)\otimes\Lambda^j(\mathfrak{n}^*)=E_j|_{K_M}. \tag{1-24}$$

Let $\mathcal{E}_j=G\times_K E_j$ be a $\mathbb{Z}_2$-graded vector bundle on X. It descends to a $\mathbb{Z}_2$-graded vector bundle $\mathcal{F}_j$ on Z. Let C_j be a Casimir operator of G action on $C^\infty(Z,\mathcal{F}_j\otimes_\mathbb{R}F)$. In Theorem 7.6, we show that there are $\sigma_j \in \mathbb{R}$ and an odd polynomial P_j such that if $\mathrm{Re}(\sigma) \gg 1$, $Z_j(\sigma)$ is holomorphic and

$$Z_j(\sigma) = \det_{\mathrm{gr}}(C_j + \sigma_j + \sigma^2) \exp(r \operatorname{vol}(Z) P_j(\sigma)), \tag{1-25}$$

where $\det_{\mathrm{gr}}$ is the zeta regularized $\mathbb{Z}_2$-graded determinant. In particular, $Z_j(\sigma)$ extends meromorphically to $\mathbb{C}$.

By a direct calculation of linear algebra, we have

$$R_\rho(\sigma) = \prod_{j=0}^{\dim \mathfrak{n}} Z_j\big(\sigma + \big(j - \tfrac{1}{2}\dim \mathfrak{n}\big)|\alpha|\big)^{(-1)^{j-1}}, \tag{1-26}$$

from which we get the meromorphic extension of $R_\rho(\sigma)$. Note that the meromorphic function

$$T(\sigma) = \prod_{i=1}^{\dim Z} \det(\sigma + \Box^Z|_{\Omega^i(Z,F)})^{(-1)^i i} \tag{1-27}$$

has a Laurent expansion near $\sigma = 0$,

$$T(\sigma) = T(F)^2 \sigma^{\chi'(X,F)} + \mathcal{O}(\sigma^{\chi'(X,F)+1}), \tag{1-28}$$

where $\chi'(X, F)$ is the derived Euler number; see (2-8). Note also that the Hodge Laplacian $\Box^Z$ coincides with the Casimir operator acting on $\Omega^\bullet(Z, F)$. The Laurent expansion (1-19) can be deduced from (1-25)–(1-28) and the identity in $\mathrm{RO}(K)$,

$$\sum_{i=1}^{\dim \mathfrak{p}} (-1)^{i-1} i \Lambda^i(\mathfrak{p}^*) = \sum_{j=0}^{\dim \mathfrak{n}} (-1)^j E_j. \tag{1-29}$$

1H. *Proof of* (1-20). To understand how the acyclicity of F is reflected in the function $R_\rho(\sigma)$, we need some deep results of representation theory. Let $\hat{p} : \Gamma\backslash G \to Z$ be the natural projection. The enveloping algebra of $U(\mathfrak{g})$ acts on $C^\infty(\Gamma\backslash G, \hat{p}^*F)$. Let $\mathcal{Z}(\mathfrak{g})$ be the center of $U(\mathfrak{g})$. Let $V^\infty \subset C^\infty(\Gamma\backslash G, \hat{p}^*F)$ be the subspace of $C^\infty(\Gamma\backslash G, \hat{p}^*F)$ on which the action of $\mathcal{Z}(\mathfrak{g})$ vanishes, and let V be the closure of V^∞ in $L^2(\Gamma\backslash G, \hat{p}^*F)$. Then V is a unitary representation of G. The compactness of $\Gamma\backslash G$ implies that V is a finite sum of irreducible unitary representations of G. By standard arguments [Borel and Wallach 2000, Chapter VII, Theorem 3.2, Corollary 3.4], the cohomology $H^\bullet(Z, F)$ is canonically isomorphic to the $(\mathfrak{g}, K)$-cohomology $H^\bullet(\mathfrak{g}, K; V)$ of V.

In [Vogan and Zuckerman 1984; Vogan 1984], the authors classified all irreducible unitary representations with nonzero $(\mathfrak{g}, K)$-cohomology. On the other hand, Salamanca-Riba [1999] showed that any irreducible unitary representation with vanishing $\mathcal{Z}(\mathfrak{g})$-action is in the class specified by Vogan and Zuckerman, which means that it possesses nonzero $(\mathfrak{g}, K)$-cohomology.

By the above considerations, the acyclicity of F is equivalent to $V = 0$. This is essentially the algebraic ingredient in the proof of (1-20). Indeed, in Corollary 8.18, we give a formula for the constants C_ρ and r_ρ, obtained by Hecht–Schmid formula [1983] with the help of the $\mathfrak{n}$-homology of V.

1I. ***The organization of the article.*** This article is organized as follows. In Section 2, we recall the definitions of certain characteristic forms and of the analytic torsion.

In Section 3, we introduce the reductive groups and the fundamental rank $\delta(G)$ of G.

In Section 4, we introduce the symmetric space. We recall basic principles for the Selberg trace formula, and we state formulas by Bismut [2011, Theorem 6.1.1] for semisimple orbital integrals. We recall the proof, given in Theorem 7.9.1 of the same paper, of a vanishing result of the analytic torsion $T(F)$ in the case $\delta(G) \neq 1$, which is originally due to Moscovici and Stanton [1991, Corollary 2.2].

In Section 5, we introduce the dynamical zeta function $R_\rho(\sigma)$, and we state Theorem 1.1 as Theorem 5.5. We prove Theorem 1.1 when $\delta(G) \neq 1$ or when G has noncompact center.

Sections 6–8 are devoted to establishing Theorem 1.1 when G has compact center and when $\delta(G) = 1$.

In Section 6, we introduce geometric objects associated with such reductive groups G.

In Section 7, we introduce Selberg zeta functions, and we prove that $R_\rho(\sigma)$ extends meromorphically, and we establish (1-19).

Finally, in Section 8, after recalling some constructions and results of representation theory, we prove that (1-20) holds.

Throughout the paper, we use the superconnection formalism of [Quillen 1985] and [Berline et al. 2004, Section 1.3]. If A is a $\mathbb{Z}_2$-graded algebra and if $a, b \in A$, the supercommutator $[a, b]$ is given by

$$[a, b] = ab - (-1)^{\deg a \deg b} ba. \tag{1-30}$$

If B is another $\mathbb{Z}_2$-graded algebra, we denote by $A \widehat{\otimes} B$ the super tensor product algebra of A and B. If $E = E^+ \oplus E^-$ is a $\mathbb{Z}_2$-graded vector space, the algebra $\mathrm{End}(E)$ is $\mathbb{Z}_2$-graded. If $\tau = \pm 1$ on $E^\pm$ and if $a \in \mathrm{End}(E)$, the supertrace $\mathrm{Tr_s}[a]$ is defined by

$$\mathrm{Tr_s}[a] = \mathrm{Tr}[\tau a]. \tag{1-31}$$

We make the convention that $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}^* = \{1, 2, \dots\}$.

2. Characteristic forms and analytic torsion

The purpose of this section is to recall some basic constructions and properties of characteristic forms and the analytic torsion.

This section is organized as follows. In Section 2A, we recall the construction of the Euler form, the $\widehat{A}$-form and the Chern character form.

In Section 2B, we introduce the regularized determinant.

Finally, in Section 2C, we recall the definition of the analytic torsion of flat vector bundles.

2A. ***Characteristic forms.*** If V is a real or complex vector space of dimension n, we denote by V^* the dual space and by $\Lambda^\bullet(V) = \sum_{i=0}^n \Lambda^i(V)$ its exterior algebra. Let Z be a smooth manifold. If V is a vector bundle on Z, we denote by $\Omega^\bullet(Z, V)$ the space of smooth differential forms with values in V. When $V = \mathbb{R}$, we write $\Omega^\bullet(Z)$ instead.

Let E be a real Euclidean vector bundle of rank m with a metric connection ∇^E. Let $R^E = \nabla^{E,2}$ be the curvature of ∇^E. It is a 2-form with values in antisymmetric endomorphisms of E.

If A is an antisymmetric matrix, denote by $\mathrm{Pf}[A]$ the Pfaffian [Bismut and Zhang 1992, Equation (3.3)] of A. Then $\mathrm{Pf}[A]$ is a polynomial function of A which is a square root of $\det[A]$. Let $o(E)$ be the orientation line of E. The Euler form $e(E, \nabla^E)$ of (E, ∇^E) is given by

$$e(E, \nabla^E) = \mathrm{Pf}\left[\frac{R^E}{2\pi}\right] \in \Omega^m(Z, o(E)). \tag{2-1}$$

If m is odd, then $e(E, \nabla^E) = 0$.

For $x \in \mathbb{C}$, set

$$\widehat{A}(x) = \frac{x/2}{\sinh(x/2)}. \tag{2-2}$$

The form $\widehat{A}(E, \nabla^E)$ of (E, ∇^E) is given by

$$\widehat{A}(E, \nabla^E) = \left[\det\left(\widehat{A}\left(-\frac{R^E}{2i\pi}\right)\right)\right]^{\frac{1}{2}} \in \Omega^\bullet(Z). \tag{2-3}$$

If E' is a complex Hermitian vector bundle equipped with a metric connection $\nabla^{E'}$ with curvature $R^{E'}$, the Chern character form $\mathrm{ch}(E', \nabla^{E'})$ of $(E', \nabla^{E'})$ is given by

$$\mathrm{ch}(E', \nabla^{E'}) = \mathrm{Tr}\left[\exp\left(-\frac{R^{E'}}{2i\pi}\right)\right] \in \Omega^\bullet(Z). \tag{2-4}$$

The differential forms $e(E, \nabla^E)$, $\widehat{A}(E, \nabla^E)$ and $\mathrm{ch}(E', \nabla^{E'})$ are closed. They are the Chern–Weil representatives of the Euler class of E, the $\widehat{A}$-genus of E and the Chern character of E'.

2B. ***Regularized determinant.*** Let (Z, g^{TZ}) be a smooth closed Riemannian manifold of dimension m. Let (E, g^E) be a Hermitian vector bundle on Z. The metrics g^{TZ}, g^E induce an L^2-metric on $C^\infty(Z, E)$.

Let P be a second-order elliptic differential operator acting on $C^\infty(Z, E)$. Suppose that P is formally self-adjoint and nonnegative. Let P^{-1} be the inverse of P acting on the orthogonal space to $\ker(P)$. For $\mathrm{Re}(s) > m/2$, set

$$\theta_P(s) = -\mathrm{Tr}[(P^{-1})^s]. \tag{2-5}$$

By [Seeley 1967] or [Berline et al. 2004, Proposition 9.35], $\theta(s)$ has a meromorphic extension to $s \in \mathbb{C}$ which is holomorphic at $s = 0$. The regularized determinant of P is defined as

$$\det(P) = \exp(\theta_P'(0)). \tag{2-6}$$

Assume now that P is formally self-adjoint and bounded from below. Denote by $\mathrm{Sp}(P)$ the spectrum of P. For $\lambda \in \mathrm{Sp}(P)$, set

$$m_P(\lambda) = \dim_{\mathbb{C}} \ker(P - \lambda) \tag{2-7}$$

to be its multiplicity. If $\sigma \in \mathbb{R}$ is such that $P + \sigma > 0$, then $\det(P + \sigma)$ is defined by (2-6). Voros [1987] has shown that the function $\sigma \to \det(P + \sigma)$, defined for $\sigma \gg 1$, extends holomorphically to $\mathbb{C}$ with zeros at $\sigma = -\lambda$ of the order $m_P(\lambda)$, where $\lambda \in \mathrm{Sp}(P)$.

2C. ***Analytic torsion.*** Let Z be a smooth connected closed manifold of dimension m with fundamental group Γ. Let F be a complex flat vector bundle on Z of rank r. Equivalently, F can be obtained via a complex representation $\rho : \Gamma \to \mathrm{GL}_r(\mathbb{C})$.

Let $H^\bullet(Z, F) = \bigoplus_{i=0}^m H^i(Z, F)$ be the cohomology of the sheaf of locally flat sections of F. We define the Euler number and the derived Euler number by

$$\chi(Z, F) = \sum_{i=0}^{m} (-1)^i \dim_{\mathbb{C}} H^i(Z, F), \quad \chi'(Z, F) = \sum_{i=1}^{m} (-1)^i i \dim_{\mathbb{C}} H^i(Z, F). \tag{2-8}$$

Let $(\Omega^\bullet(Z, F), d^Z)$ be the de Rham complex of smooth sections of $\Lambda^\bullet(T^*Z) \otimes_{\mathbb{R}} F$ on Z. We have the canonical isomorphism of vector spaces

$$H^\bullet(\Omega^\bullet(Z, F), d^Z) \simeq H^\bullet(Z, F). \tag{2-9}$$

In the sequel, we will also consider the trivial line bundle $\mathbb{R}$. We denote simply by $H^\bullet(Z)$ and $\chi(Z)$ the corresponding objects. Note that, in this case, the complex dimension in (2-8) should be replaced by the real dimension.

Let g^{TZ} be a Riemannian metric on TZ, and let g^F be a Hermitian metric on F. They induce an L^2-metric $\langle \cdot, \cdot \rangle_{\Omega^\bullet(Z,F)}$ on $\Omega^\bullet(Z, F)$. Let $d^{Z,*}$ be the formal adjoint of d^Z with respect to $\langle \cdot, \cdot \rangle_{\Omega^\bullet(Z,F)}$. Put

$$D^Z = d^Z + d^{Z,*}, \quad \square^Z = D^{Z,2} = [d^Z, d^{Z,*}]. \tag{2-10}$$

Then, $\square^Z$ is a formally self-adjoint nonnegative second-order elliptic operator acting on $\Omega^\bullet(Z, F)$. By Hodge theory, we have the canonical isomorphism of vector spaces

$$\ker \square^Z \simeq H^\bullet(Z, F). \tag{2-11}$$

Definition 2.1. The analytic torsion of F is a positive real number defined by

$$T(F, g^{TZ}, g^F) = \prod_{i=1}^{m} \det(\square^Z|_{\Omega^i(Z,F)})^{(-1)^i i/2}. \tag{2-12}$$

Recall that the flat vector bundle F carries a flat metric g^F if and only if the holonomy representation ρ factors through $\mathrm{U}(r)$. In this case, F is said to be unitarily flat. If Z is an even-dimensional orientable manifold and if F is unitarily flat with a flat metric g^F, by Poincaré duality, $T(F, g^{TZ}, g^F) = 1$. If $\dim Z$ is odd and if $H^\bullet(Z, F) = 0$, by [Bismut and Zhang 1992, Theorem 4.7], then $T(F, g^{TZ}, g^F)$ does not depend on g^{TZ} or g^F. In the sequel, we write instead

$$T(F) = T(F, g^{TZ}, g^F). \tag{2-13}$$

By Section 2B,

$$T(\sigma) = \prod_{i=1}^{\dim Z} \det(\sigma + \square^Z|_{\Omega^i(Z,F)})^{(-1)^i i} \tag{2-14}$$

is meromorphic on $\mathbb{C}$. When $\sigma \to 0$, we have

$$T(\sigma) = T(F)^2 \sigma^{\chi'(Z,F)} + \mathcal{O}(\sigma^{\chi'(Z,F)+1}). \tag{2-15}$$

3. Preliminaries on reductive groups

The purpose of this section is to recall some basic facts about reductive groups.

This section is organized as follows. In Section 3A, we introduce the reductive group G.

In Section 3B, we introduce the semisimple elements of G, and we recall some related constructions.

In Section 3C, we recall some properties of Cartan subgroups of G. We introduce a nonnegative integer $\delta(G)$, which has paramount importance in the whole article. We also recall Weyl's integral formula on reductive groups.

Finally, in Section 3D, we introduce the regular elements of G.

3A. ***The reductive group.*** Let G be a linear connected real reductive group [Knapp 1986, p. 3]; that is, G is a closed connected group of real matrices that is stable under transpose. Let $\theta \in \mathrm{Aut}(G)$ be the Cartan involution. Let K be the maximal compact subgroup of G of the points of G that are fixed by θ.

Let $\mathfrak{g}$ be the Lie algebra of G, and let $\mathfrak{k} \subset \mathfrak{g}$ be the Lie algebra of K. The Cartan involution θ acts naturally as a Lie algebra automorphism of $\mathfrak{g}$. Then $\mathfrak{k}$ is the eigenspace of θ associated with the eigenvalue 1. Let $\mathfrak{p}$ be the eigenspace with the eigenvalue -1, so that

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}. \tag{3-1}$$

Then we have

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}. \tag{3-2}$$

Put

$$m = \dim \mathfrak{p}, \quad n = \dim \mathfrak{k}. \tag{3-3}$$

By [Knapp 1986, Proposition 1.2], we have the diffeomorphism

$$(Y, k) \in \mathfrak{p} \times K \to e^{Y} k \in G. \tag{3-4}$$

Let B be a real-valued nondegenerate bilinear symmetric form on $\mathfrak{g}$ which is invariant under the adjoint action Ad of G on $\mathfrak{g}$, and also under θ. Then (3-1) is an orthogonal splitting of $\mathfrak{g}$ with respect to B. We assume B to be positive on $\mathfrak{p}$ and negative on $\mathfrak{k}$. The form $\langle \cdot\,, \cdot \rangle = -B(\cdot\,, \theta\, \cdot)$ defines an $\mathrm{Ad}(K)$-invariant scalar product on $\mathfrak{g}$ such that the splitting (3-1) is still orthogonal. We denote by $|\cdot|$ the corresponding norm.

Let $Z_G \subset G$ be the center of G with Lie algebra $\mathfrak{z}_{\mathfrak{g}} \subset \mathfrak{g}$. Set

$$\mathfrak{z}_{\mathfrak{p}} = \mathfrak{z}_{\mathfrak{g}} \cap \mathfrak{p}, \quad \mathfrak{z}_{\mathfrak{k}} = \mathfrak{z}_{\mathfrak{g}} \cap \mathfrak{k}. \tag{3-5}$$

By [Knapp 1986, Corollary 1.3], Z_G is reductive. As in (3-1) and (3-4), we have the Cartan decomposition

$$\mathfrak{z}_{\mathfrak{g}} = \mathfrak{z}_{\mathfrak{p}} \oplus \mathfrak{z}_{\mathfrak{k}}, \quad Z_G = \exp(\mathfrak{z}_{\mathfrak{p}})(Z_G \cap K). \tag{3-6}$$

Let $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ be the complexification of $\mathfrak{g}$ and let $\mathfrak{u} = \sqrt{-1}\mathfrak{p} \oplus \mathfrak{k}$ be the compact form of $\mathfrak{g}$. Let $G_{\mathbb{C}}$ and U be the connected group of complex matrices associated with the Lie algebras $\mathfrak{g}_{\mathbb{C}}$ and $\mathfrak{u}$. By [Knapp 1986, Propositions 5.3 and 5.6], if G has compact center, i.e., its center Z_G is compact, then $G_{\mathbb{C}}$ is a linear connected complex reductive group with maximal compact subgroup U.

Let $U(\mathfrak{g})$ be the enveloping algebra of $\mathfrak{g}$, and let $\mathcal{Z}(\mathfrak{g}) \subset U(\mathfrak{g})$ be the center of $U(\mathfrak{g})$. Let $C^{\mathfrak{g}} \in U(\mathfrak{g})$ be the Casimir element. If $e_1, \dots, e_m$ is an orthonormal basis of $\mathfrak{p}$, and if $e_{m+1}, \dots, e_{m+n}$ is an orthonormal basis of $\mathfrak{k}$, then

$$C^{\mathfrak{g}} = -\sum_{i=1}^{m} e_i^2 + \sum_{i=m+1}^{n+m} e_i^2. \tag{3-7}$$

Classically, $C^{\mathfrak{g}} \in \mathcal{Z}(\mathfrak{g})$.

We define $C^{\mathfrak{k}}$ similarly. Let τ be a finite-dimensional representation of K on V. We denote by $C^{\mathfrak{k},V}$ or $C^{\mathfrak{k},\tau} \in \operatorname{End}(V)$ the corresponding Casimir operator acting on V, so that

$$C^{\mathfrak{k},V} = C^{\mathfrak{k},\tau} = \sum_{i=m+1}^{m+n} \tau(e_i)^2. \tag{3-8}$$

3B. ***Semisimple elements.*** If $\gamma \in G$, we denote by $Z(\gamma) \subset G$ the centralizer of γ in G, and by $\mathfrak{z}(\gamma) \subset \mathfrak{g}$ its Lie algebra. If $a \in \mathfrak{g}$, let $Z(a) \subset G$ be the stabilizer of a in G, and let $\mathfrak{z}(a) \subset \mathfrak{g}$ be its Lie algebra.

An element $\gamma \in G$ is said to be semisimple if γ can be conjugated to $e^a k^{-1}$ such that

$$a \in \mathfrak{p}, \quad k \in K, \quad \operatorname{Ad}(k)a = a. \tag{3-9}$$

Let $\gamma = e^a k^{-1}$ be such that (3-9) holds. By [Bismut 2011, Equations (3.3.4), (3.3.6)], we have

$$Z(\gamma) = Z(a) \cap Z(k), \quad \mathfrak{z}(\gamma) = \mathfrak{z}(a) \cap \mathfrak{z}(k). \tag{3-10}$$

Set

$$\mathfrak{p}(\gamma) = \mathfrak{z}(\gamma) \cap \mathfrak{p}, \quad \mathfrak{k}(\gamma) = \mathfrak{z}(\gamma) \cap \mathfrak{k}. \tag{3-11}$$

From (3-10) and (3-11), we get

$$\mathfrak{z}(\gamma) = \mathfrak{p}(\gamma) \oplus \mathfrak{k}(\gamma). \tag{3-12}$$

By [Knapp 2002, Proposition 7.25], $Z(\gamma)$ is a reductive subgroup of G with maximal compact subgroup $K(\gamma) = Z(\gamma) \cap K$, and with Cartan decomposition (3-12). Let $Z^0(\gamma)$ be the connected component of the identity in $Z(\gamma)$. Then $Z^0(\gamma)$ is a reductive subgroup of G, with maximal compact subgroup $Z^0(\gamma) \cap K$. Also, $Z^0(\gamma) \cap K$ coincides with $K^0(\gamma)$, the connected component of the identity in $K(\gamma)$.

An element $\gamma \in G$ is said to be elliptic if γ is conjugated to an element of K. Let $\gamma \in G$ be semisimple and nonelliptic. Up to conjugation, we can assume $\gamma = e^a k^{-1}$ such that (3-9) holds and that $a \neq 0$. By (3-10), $a \in \mathfrak{p}(\gamma)$. Let $\mathfrak{z}^{a,\perp}(\gamma)$, $\mathfrak{p}^{a,\perp}(\gamma)$ be respectively the orthogonal spaces to a in $\mathfrak{z}(\gamma)$, $\mathfrak{p}(\gamma)$, so that

$$\mathfrak{z}^{a,\perp}(\gamma) = \mathfrak{p}^{a,\perp}(\gamma) \oplus \mathfrak{k}(\gamma). \tag{3-13}$$

Moreover, $\mathfrak{z}^{a,\perp}(\gamma)$ is a Lie algebra. Let $Z^{a,\perp,0}(\gamma)$ be the connected subgroup of $Z^0(\gamma)$ that is associated with the Lie algebra $\mathfrak{z}^{a,\perp}(\gamma)$. By [Bismut 2011, Equation (3.3.11)], $Z^{a,\perp,0}(\gamma)$ is reductive with maximal compact subgroup $K^0(\gamma)$ with Cartan decomposition (3-13), and

$$Z^0(\gamma) = \mathbb{R} \times Z^{a,\perp,0}(\gamma), \tag{3-14}$$

so that e^{ta} maps into $t|a|$.

3C. ***Cartan subgroups.*** A Cartan subalgebra of $\mathfrak{g}$ is a maximal abelian subalgebra of $\mathfrak{g}$. A Cartan subgroup of G is the centralizer of a Cartan subalgebra.

By [Knapp 1986, Theorem 5.22], there is only a finite number of nonconjugate (via K) θ-stable Cartan subalgebras $\mathfrak{h}_1, \dots, \mathfrak{h}_{l_0}$. Let $H_1, \dots, H_{l_0}$ be the corresponding Cartan subgroups. Clearly, the Lie algebra of H_i is $\mathfrak{h}_i$. Set

$$\mathfrak{h}_{i\mathfrak{p}} = \mathfrak{h}_i \cap \mathfrak{p}, \quad \mathfrak{h}_{i\mathfrak{k}} = \mathfrak{h}_i \cap \mathfrak{k}. \tag{3-15}$$

We call $\dim \mathfrak{h}_{i\mathfrak{p}}$ the noncompact dimension of $\mathfrak{h}_i$. By [Knapp 1986, Theorem 5.22(c); 2002, Proposition 7.25], H_i is an abelian reductive group with maximal compact subgroup $H_i \cap K$, and with Cartan decomposition

$$\mathfrak{h}_i = \mathfrak{h}_{i\mathfrak{p}} \oplus \mathfrak{h}_{i\mathfrak{k}}, \quad H_i = \exp(\mathfrak{h}_{i\mathfrak{p}})(H_i \cap K). \tag{3-16}$$

Note that in general, H_i is not necessarily connected.

Let $W(H_i, G)$ be the Weyl group. If $N_K(\mathfrak{h}_i) \subset K$ and $Z_K(\mathfrak{h}_i) \subset K$ are the normalizer and centralizer of $\mathfrak{h}_i$ in K, then

$$W(H_i, G) = N_K(\mathfrak{h}_i)/Z_K(\mathfrak{h}_i). \tag{3-17}$$

Throughout, we fix a maximal torus T of K. Let $\mathfrak{t} \subset \mathfrak{k}$ be the Lie algebra of T. Set

$$\mathfrak{b} = \{Y \in \mathfrak{p} : [Y, \mathfrak{t}] = 0\}. \tag{3-18}$$

By (3-5) and (3-18), we have

$$\mathfrak{z}_{\mathfrak{p}} \subset \mathfrak{b}. \tag{3-19}$$

Put

$$\mathfrak{h} = \mathfrak{b} \oplus \mathfrak{t}. \tag{3-20}$$

By [Knapp 1986, Theorem 5.22(b)], $\mathfrak{h}$ is the θ-stable Cartan subalgebra of $\mathfrak{g}$ with minimal noncompact dimension. Also, every θ-stable Cartan subalgebra with minimal noncompact dimension is conjugated to $\mathfrak{h}$ by an element of K. Moreover, the corresponding Cartan subgroup $H \subset G$ of G is connected, so that

$$H = \exp(\mathfrak{b})T. \tag{3-21}$$

We may assume that $\mathfrak{h}_1 = \mathfrak{h}$ and $H_1 = H$.

Note that the complexification $\mathfrak{h}_{i\mathbb{C}} = \mathfrak{h}_i \otimes_{\mathbb{R}} \mathbb{C}$ of $\mathfrak{h}_i$ is a Cartan subalgebra of $\mathfrak{g}_{\mathbb{C}}$. All the $\mathfrak{h}_{i\mathbb{C}}$ are conjugated by inner automorphisms of $\mathfrak{g}_{\mathbb{C}}$. Their common complex dimension $\dim_{\mathbb{C}} \mathfrak{h}_{i\mathbb{C}}$ is called the complex rank $\mathrm{rk}_{\mathbb{C}}(G)$ of G.

Definition 3.1. Put

$$\delta(G) = \mathrm{rk}_{\mathbb{C}}(G) - \mathrm{rk}_{\mathbb{C}}(K) \in \mathbb{N}. \tag{3-22}$$

By (3-18) and (3-22), we have

$$\delta(G) = \dim \mathfrak{b}. \tag{3-23}$$

Note that $m - \delta(G)$ is even. We will see that $\delta(G)$ plays an important role in our article.

Remark 3.2. If $\mathfrak{g}$ is a real reductive Lie algebra, then $\delta(\mathfrak{g}) \in \mathbb{N}$ can be defined in the same way as in (3-23). Since $\mathfrak{g}$ is reductive, by [Knapp 2002, Corollary 1.56], we have

$$\mathfrak{g} = \mathfrak{z}_{\mathfrak{g}} \oplus [\mathfrak{g}, \mathfrak{g}], \tag{3-24}$$

where $[\mathfrak{g}, \mathfrak{g}]$ is a semisimple Lie algebra. By (3-6) and (3-24), we have

$$\delta(\mathfrak{g}) = \dim \mathfrak{z}_{\mathfrak{p}} + \delta([\mathfrak{g}, \mathfrak{g}]). \tag{3-25}$$

Proposition 3.3. *The element $\gamma \in G$ is semisimple if and only if γ can be conjugated into $\bigcup_{i=1}^{l_0} H_i$. In this case,*

$$\delta(G) \leqslant \delta(Z^0(\gamma)). \tag{3-26}$$

The two sides of (3-26) *are equal if and only if γ can be conjugated into H.*

Proof. If $\gamma \in H_i$, by the Cartan decomposition (3-16), there exist $a \in \mathfrak{h}_{i\mathfrak{p}}$ and $k \in K \cap H_i$ such that $\gamma = e^a k^{-1}$. Since H_i is the centralizer of $\mathfrak{h}_i$, we have $\mathrm{Ad}(\gamma)a = a$. Therefore, $\mathrm{Ad}(k)a = a$, so that γ is semisimple.

Assume that $\gamma \in G$ is semisimple and is such that (3-9) holds. We claim that

$$\mathrm{rk}_{\mathbb{C}}(G) = \mathrm{rk}_{\mathbb{C}}(Z^0(\gamma)). \tag{3-27}$$

Indeed, let $\mathfrak{h}' \subset \mathfrak{g}$ be a θ-invariant Cartan subalgebra of $\mathfrak{g}$ containing a. Then, $\mathfrak{h}' \subset \mathfrak{z}(a)$. It implies

$$\mathrm{rk}_{\mathbb{C}}(G) = \mathrm{rk}_{\mathbb{C}}(Z^0(a)). \tag{3-28}$$

By choosing a maximal torus T containing k, by (3-20), we have $\mathfrak{h} \subset \mathfrak{z}(k)$. Then

$$\mathrm{rk}_{\mathbb{C}}(G) = \mathrm{rk}_{\mathbb{C}}(Z^0(k)). \tag{3-29}$$

If we replace G by $Z^0(a)$ in (3-29), by (3-10), we get

$$\mathrm{rk}_{\mathbb{C}}(Z^0(a)) = \mathrm{rk}_{\mathbb{C}}(Z^0(\gamma)). \tag{3-30}$$

By (3-28) and (3-30), we get (3-27).

Let $\mathfrak{h}(\gamma) \subset \mathfrak{z}(\gamma)$ be the θ-invariant Cartan subalgebra defined as in (3-20) when G is replaced by $Z^0(\gamma)$. By (3-27), $\mathfrak{h}(\gamma)$ is also a Cartan subalgebra of $\mathfrak{g}$. Moreover, γ is an element of the Cartan subgroup of G associated to $\mathfrak{h}(\gamma)$. In particular, γ can be conjugated into some H_i.

By the minimality of noncompact dimension of $\mathfrak{h}$, we have

$$\delta(G) = \dim \mathfrak{h} \cap \mathfrak{p} \leqslant \dim \mathfrak{h}(\gamma) \cap \mathfrak{p} = \delta(Z^0(\gamma)), \tag{3-31}$$

which completes the proof of (3-26).

It is obvious that if γ can be conjugated into H, the equality in (3-31) holds. If the equality holds in (3-31), by the uniqueness of the Cartan subalgebra with minimal noncompact dimension, there is $k' \in K$ such that

$$\mathrm{Ad}(k')\mathfrak{h}(\gamma) = \mathfrak{h}, \tag{3-32}$$

which implies that $k'\gamma k'^{,-1} \in H$. □

Now we recall the Weyl integral formula on G, which will be used in Section 8. Let dv_{H_i} and $dv_{H_i\backslash G}$ be respectively the Riemannian volumes on H_i and $H_i\backslash G$ induced by $-B(\cdot,\theta\cdot)$. By [Knapp 2002, Theorem 8.64], for a nonnegative measurable function f on G, we have

$$\int_{g\in G} f(g)\,dv_G=\sum_{i=1}^{l_0}\frac{1}{|W(H_i,G)|}\int_{\gamma\in H_i}\left(\int_{g\in H_i\backslash G} f(g^{-1}\gamma g)\,dv_{H_i\backslash G}\right)\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\right|dv_{H_i}. \tag{3-33}$$

3D. ***Regular elements.*** For $0\leqslant j\leqslant m+n-\mathrm{rk}_{\mathbb{C}}(G)$, let D_j be the analytic function on G such that, for $t\in\mathbb{R}$ and $\gamma\in G$, we have

$$\det(t+1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}}=t^{\mathrm{rk}_{\mathbb{C}}(G)}\left(\sum_{j=0}^{m+n-\mathrm{rk}_{\mathbb{C}}(G)} D_j(\gamma)t^j\right). \tag{3-34}$$

If $\gamma\in H_i$, then

$$D_0(\gamma)=\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}. \tag{3-35}$$

We call $\gamma\in G$ regular if $D_0(\gamma)\neq 0$. Let $G'\subset G$ be the subset of regular elements of G. Then G' is open in G such that $G-G'$ has zero measure with respect to the Riemannian volume dv_G on G induced by $-B(\cdot,\theta\cdot)$. For $1\leqslant i\leqslant l_0$, set

$$H_i'=H_i\cap G',\quad G_i'=\bigcup_{g\in G} g^{-1}H_i'g. \tag{3-36}$$

By [Knapp 1986, Theorem 5.22(d)], G_i' is open, and we have the disjoint union

$$G'=\coprod_{1\leqslant i\leqslant l_0} G_i'. \tag{3-37}$$

4. Orbital integrals and Selberg trace formula

The purpose of this section is to recall the semisimple orbital integral formula of [Bismut 2011, Theorem 6.1.1] and the Selberg trace formula.

This section is organized as follows. In Section 4A, we introduce the Riemannian symmetric space $X=G/K$, and we give a formula for its Euler form.

In Section 4B, we recall the definition of semisimple orbital integrals.

In Section 4C, we recall Bismut's explicit formula for the semisimple orbital integrals associated to the heat operator of the Casimir element.

In Section 4D, we introduce a discrete torsion-free cocompact subgroup Γ of G. We state the Selberg trace formula.

Finally, in Section 4E, we recall Bismut's proof of a vanishing result on the analytic torsion in the case $\delta(G)\neq 1$, which is originally due to Moscovici and Stanton [1991].

4A. ***The symmetric space.*** We use the notation of Section 3. Let $\omega^{\mathfrak{g}}$ be the canonical left-invariant 1-form on G with values in $\mathfrak{g}$, and let $\omega^{\mathfrak{p}},\omega^{\mathfrak{k}}$ be its components in $\mathfrak{p},\mathfrak{k}$, so that

$$\omega^{\mathfrak{g}}=\omega^{\mathfrak{p}}+\omega^{\mathfrak{k}}. \tag{4-1}$$

Let $X = G/K$ be the associated symmetric space. Then

$$p : G \to X = G/K \tag{4-2}$$

is a K-principle bundle, equipped with the connection form $\omega^{\mathfrak{k}}$. By (3-2) and (4-1), the curvature of $\omega^{\mathfrak{k}}$ is given by

$$\Omega^{\mathfrak{k}} = -\tfrac{1}{2}[\omega^{\mathfrak{p}}, \omega^{\mathfrak{p}}]. \tag{4-3}$$

Let τ be a finite-dimensional orthogonal representation of K on the real Euclidean space E_τ. Then $\mathcal{E}_\tau = G \times_K E_\tau$ is a real Euclidean vector bundle on X, which is naturally equipped with a Euclidean connection $\nabla^{\mathcal{E}_\tau}$. The space of smooth sections $C^\infty(X, \mathcal{E}_\tau)$ on X can be identified with the K-invariant subspace $C^\infty(G, E_\tau)^K$ of smooth E_τ-valued functions on G. Let $C^{\mathfrak{g},X,\tau}$ be the Casimir element of G acting on $C^\infty(X, \mathcal{E}_\tau)$. Then $C^{\mathfrak{g},X,\tau}$ is a formally self-adjoint second-order elliptic differential operator which is bounded from below.

Observe that K acts isometrically on $\mathfrak{p}$. Using the above construction, the tangent bundle $TX = G \times_K \mathfrak{p}$ is equipped with a Euclidean metric g^{TX} and a Euclidean connection ∇^{TX}. Also, ∇^{TX} is the Levi-Civita connection on (TX, g^{TX}) with curvature R^{TX}. Classically, (X, g^{TX}) is a Riemannian manifold of nonpositive sectional curvature. For $x, y \in X$, we denote by $d_X(x, y)$ the Riemannian distance on X.

If $E_\tau = \Lambda^\bullet(\mathfrak{p}^*)$, then $C^\infty(X, \mathcal{E}_\tau) = \Omega^\bullet(X)$. In this case, we write $C^{\mathfrak{g},X} = C^{\mathfrak{g},X,\tau}$. By [Bismut 2011, Proposition 7.8.1], $C^{\mathfrak{g},X}$ coincides with the Hodge Laplacian acting on $\Omega^\bullet(X)$.

Let us state a formula for $e(TX, \nabla^{TX})$. Let $o(TX)$ be the orientation line of TX. Let dv_X be the G-invariant Riemannian volume form on X. If $\alpha \in \Omega^\bullet(X, o(TX))$ is of maximal degree and G-invariant, set $[\alpha]^{\max} \in \mathbb{R}$ such that

$$\alpha = [\alpha]^{\max}\, dv_X. \tag{4-4}$$

Recall that if G has compact center, then U is the compact form of G. If $\delta(G) = 0$, by (3-25), G has compact center. In this case, T is a maximal torus of both U and K. Let $W(T, U)$, $W(T, K)$ be the respective the Weyl groups. Let $\mathrm{vol}(U/K)$ be the volume of U/K with respect to the volume form induced by $-B$.

Proposition 4.1. *If $\delta(G) \neq 0$, then $[e(TX, \nabla^{TX})]^{\max} = 0$. If $\delta(G) = 0$, then*

$$[e(TX, \nabla^{TX})]^{\max} = (-1)^{\frac{m}{2}} \frac{|W(T, U)|/|W(T, K)|}{\mathrm{vol}(U/K)}. \tag{4-5}$$

Proof. If G has noncompact center (thus $\delta(G) \neq 0$), it is trivial that $[e(TX, \nabla^{TX})]^{\max} = 0$. Assume now, G has compact center. By Hirzebruch proportionality [1966] (see Theorem 22.3.1 of that paper for a proof for Hermitian symmetric spaces; the proof for general case is identical), we have

$$[e(TX, \nabla^{TX})]^{\max} = (-1)^{\frac{m}{2}} \frac{\chi(U/K)}{\mathrm{vol}(U/K)}. \tag{4-6}$$

Proposition 4.1 is a consequence of (4-6), Bott's formula [1965, p. 175], Theorem II of the same paper and of the fact that $\delta(G) = \mathrm{rk}_{\mathbb{C}}(U) - \mathrm{rk}_{\mathbb{C}}(K)$. □

Let $\gamma \in G$ be a semisimple element as in (3-9). Let

$$X(\gamma) = Z(\gamma)/K(\gamma) \tag{4-7}$$

be the associated symmetric space. Clearly,

$$X(\gamma) = Z^0(\gamma)/K^0(\gamma). \tag{4-8}$$

Suppose that γ is nonelliptic. Set

$$X^{a,\perp}(\gamma) = Z^{a,\perp,0}(\gamma)/K^0(\gamma). \tag{4-9}$$

By (3-14), (4-8) and (4-9), we have

$$X(\gamma) = \mathbb{R} \times X^{a,\perp}(\gamma), \tag{4-10}$$

so that the action e^{ta} on $X(\gamma)$ is just the translation by $t|a|$ on $\mathbb{R}$.

4B. ***The semisimple orbital integrals.*** Recall that τ is a finite-dimensional orthogonal representation of K on the real Euclidean space E_τ, and that $C^{\mathfrak{g},X,\tau}$ is the Casimir element of G acting on $C^\infty(X, \mathcal{E}_\tau)$.

Let $p_t^{X,\tau}(x,x')$ be the smooth kernel of $\exp(-tC^{\mathfrak{g},X,\tau}/2)$ with respect to the Riemannian volume dv_X on X. Classically, for $t>0$, there exist $c>0$ and $C>0$ such that, for $x, x' \in X$,

$$|p_t^{X,\tau}(x,x')| \leqslant C \exp(-c\, d_X^2(x,x')). \tag{4-11}$$

Set

$$p_t^{X,\tau}(g) = p_t^{X,\tau}(p1, pg). \tag{4-12}$$

For $g \in G$ and $k, k' \in K$, we have

$$p_t^{X,\tau}(kgk') = \tau(k) p_t^{X,\tau}(g) \tau(k'). \tag{4-13}$$

Also, we can recover $p_t^{X,\tau}(x,x')$ by

$$p_t^{X,\tau}(x,x') = p_t^{X,\tau}(g^{-1}g'), \tag{4-14}$$

where $g, g' \in G$ are such that $pg = x,\ pg' = x'$.

In the sequel, we do not distinguish $p_t^{X,\tau}(x,x')$ and $p_t^{X,\tau}(g)$. We refer to both of them as being the smooth kernel of $\exp(-tC^{\mathfrak{g},X,\tau}/2)$.

Let $dv_{K^0(\gamma)\backslash K}$ and $dv_{Z^0(\gamma)\backslash G}$ be the Riemannian volumes on $K^0(\gamma)\backslash K$ and $Z^0(\gamma)\backslash G$ induced by $-B(\cdot, \theta\,\cdot)$. Let $\mathrm{vol}(K^0(\gamma)\backslash K)$ be the volume of $K^0(\gamma)\backslash K$ with respect to $dv_{K^0(\gamma)\backslash K}$.

Definition 4.2. Let $\gamma \in G$ be semisimple. The orbital integral of $\exp(-tC^{\mathfrak{g},X,\tau}/2)$ is defined by

$$\mathrm{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)] = \frac{1}{\mathrm{vol}(K^0(\gamma)\backslash K)} \int_{g \in Z^0(\gamma)\backslash G} \mathrm{Tr}^{E_\tau}[p_t^{X,\tau}(g^{-1}\gamma g)]\, dv_{Z^0(\gamma)\backslash G}. \tag{4-15}$$

Remark 4.3. Definition 4.2 is equivalent to [Bismut 2011, Definition 4.2.2], where the volume forms are normalized such that $\mathrm{vol}(K^0(\gamma)\backslash K) = 1$.

Remark 4.4. As the notation $\mathrm{Tr}^{[\gamma]}$ indicates, the orbital integral only depends on the conjugacy class of γ in G. However, the notation $[\gamma]$ (see Section 4D) will be used later for the conjugacy class in the discrete group Γ.

Remark 4.5. We will also consider the case where E_τ is a $\mathbb{Z}_2$-graded or virtual representation of K. We will use the notation $\mathrm{Tr_s}^{[\gamma]}[q]$ when the trace on the right-hand side of (4-15) is replaced by the supertrace on E_τ.

4C. ***Bismut's formula for semisimple orbital integrals.*** Let us first recall the explicit formula for $\mathrm{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]$ for any semisimple element $\gamma \in G$, obtained by Bismut [2011, Theorem 6.1.1].

Let $\gamma = e^a k^{-1} \in G$ be semisimple as in (3-9). Set

$$\mathfrak{z}_0 = \mathfrak{z}(a), \quad \mathfrak{p}_0 = \mathfrak{z}(a) \cap \mathfrak{p}, \quad \mathfrak{k}_0 = \mathfrak{z}(a) \cap \mathfrak{k}. \tag{4-16}$$

Then

$$\mathfrak{z}_0 = \mathfrak{p}_0 \oplus \mathfrak{k}_0. \tag{4-17}$$

By (3-10), (3-11) and (4-16), we have $\mathfrak{p}(\gamma) \subset \mathfrak{p}_0$ and $\mathfrak{k}(\gamma) \subset \mathfrak{k}_0$. Let $\mathfrak{p}_0^\perp(\gamma)$, $\mathfrak{k}_0^\perp(\gamma)$, $\mathfrak{z}_0^\perp(\gamma)$ be the orthogonal spaces of $\mathfrak{p}(\gamma)$, $\mathfrak{k}(\gamma)$, $\mathfrak{z}(\gamma)$ in $\mathfrak{p}_0$, $\mathfrak{k}_0$, $\mathfrak{z}_0$. Let $\mathfrak{p}_0^\perp$, $\mathfrak{k}_0^\perp$, $\mathfrak{z}_0^\perp$ be the orthogonal spaces of $\mathfrak{p}_0$, $\mathfrak{k}_0$, $\mathfrak{z}_0$ in $\mathfrak{p}$, $\mathfrak{k}$, $\mathfrak{z}$. Then we have

$$\mathfrak{p} = \mathfrak{p}(\gamma) \oplus \mathfrak{p}_0^\perp(\gamma) \oplus \mathfrak{p}_0^\perp, \quad \mathfrak{k} = \mathfrak{k}(\gamma) \oplus \mathfrak{k}_0^\perp(\gamma) \oplus \mathfrak{k}_0^\perp. \tag{4-18}$$

Recall that $\widehat{A}$ is the function defined in (2-2).

Definition 4.6. For $Y \in \mathfrak{k}(\gamma)$, put

$$\begin{aligned} J_\gamma(Y) = {}& \frac{1}{\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}}} \frac{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{p}(\gamma)})}{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{k}(\gamma)})} \\ & \times \left[\frac{1}{\det(1-\mathrm{Ad}(k^{-1}))|_{\mathfrak{z}_0^\perp(\gamma)}} \frac{\det(1-\exp(-i\,\mathrm{ad}(Y))\,\mathrm{Ad}(k^{-1}))|_{\mathfrak{k}_0^\perp(\gamma)}}{\det(1-\exp(-i\,\mathrm{ad}(Y))\,\mathrm{Ad}(k^{-1}))|_{\mathfrak{p}_0^\perp(\gamma)}}\right]^{\frac{1}{2}}. \end{aligned} \tag{4-19}$$

As explained in [Bismut 2011, Section 5.5], there is a natural choice for the square root in (4-19). Moreover, J_γ is an $\mathrm{Ad}(K^0(\gamma))$-invariant analytic function on $\mathfrak{k}(\gamma)$, and there exist $c_\gamma > 0$, $C_\gamma > 0$, such that, for $Y \in \mathfrak{k}(\gamma)$,

$$|J_\gamma(Y)| \leqslant C_\gamma \exp(c_\gamma |Y|). \tag{4-20}$$

By (4-19), we have

$$J_1(Y) = \frac{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{p}})}{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{k}})}. \tag{4-21}$$

For $Y \in \mathfrak{k}(\gamma)$, let dY be the Lebesgue measure on $\mathfrak{k}(\gamma)$ induced by $-B$. Recall that $C^{\mathfrak{k},\mathfrak{p}}$ and $C^{\mathfrak{k},\mathfrak{k}}$ are defined in (3-8). The main result of [Bismut 2011, Theorem 6.1.1] is the following.

Theorem 4.7. *For $t > 0$, we have*

$$\operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)] = \frac{1}{(2\pi t)^{\dim \mathfrak{z}(\gamma)/2}} \exp\left(-\frac{|a|^2}{2t} + \frac{t}{16}\operatorname{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\operatorname{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right) \\ \times \int_{Y\in\mathfrak{k}(\gamma)} J_\gamma(Y)\operatorname{Tr}^{E_\tau}[\tau(k^{-1})\exp(-i\tau(Y))]\exp(-|Y|^2/(2t))\,dY. \tag{4-22}$$

4D. ***A discrete subgroup of G.*** Let $\Gamma \subset G$ be a discrete torsion-free cocompact subgroup of G. By [Selberg 1960, Lemma 1], Γ contains the identity element and nonelliptic semisimple elements. Also, Γ acts isometrically on the left on X. This action lifts to all the homogeneous Euclidean vector bundles $\mathcal{E}_\tau$ constructed in Section 4A, and preserves the corresponding connections.

Take $Z = \Gamma\backslash X = \Gamma\backslash G/K$. Then Z is a connected closed orientable Riemannian locally symmetric manifold with nonpositive sectional curvature. Since X is contractible, $\pi_1(Z) = \Gamma$ and X is the universal cover of Z. We denote by $\hat{p} : \Gamma\backslash G \to Z$ and $\hat{\pi} : X \to Z$ the natural projections, so that the diagram

$$\begin{tikzcd} G \arrow[r] \arrow[d, "p"] & \Gamma\backslash G \arrow[d, "\hat{p}"] \\ X \arrow[r, "\hat{\pi}"] & Z \end{tikzcd} \tag{4-23}$$

commutes.

The Euclidean vector bundle $\mathcal{E}_\tau$ descends to a Euclidean vector bundle $\mathcal{F}_\tau = \Gamma\backslash\mathcal{E}_\tau$ on Z. Take $r \in \mathbb{N}^*$. Let $\rho : \Gamma \to \mathrm{U}(r)$ be a unitary representation of Γ. Let (F, ∇^F, g^F) be the unitarily flat vector bundle on Z associated to ρ. Let $C^{\mathfrak{g},Z,\tau,\rho}$ be the Casimir element of G acting on $C^\infty(Z, \mathcal{F}_\tau \otimes_{\mathbb{R}} F)$. As in Section 4A, when $E_\tau = \Lambda^\bullet(\mathfrak{p}^*)$, we write $C^{\mathfrak{g},Z,\rho} = C^{\mathfrak{g},Z,\tau,\rho}$. Then,

$$\square^Z = C^{\mathfrak{g},Z,\rho}. \tag{4-24}$$

Recall that $p_t^{X,\tau}(x,x')$ is the smooth kernel of $\exp(-tC^{\mathfrak{g},X,\tau}/2)$ with respect to dv_X.

Proposition 4.8. *There exist $c > 0$, $C > 0$ such that, for $t > 0$ and $x \in X$, we have*

$$\sum_{\gamma\in\Gamma-\{1\}} |p_t^{X,\tau}(x,\gamma x)| \leqslant C\exp\left(-\frac{c}{t} + Ct\right). \tag{4-25}$$

Proof. By [Milnor 1968a, Remark p. 1, Lemma 2] or [Ma and Marinescu 2015, Equation (3.19)], there is $C > 0$ such that, for all $r \geqslant 0$, $x \in X$, we have

$$\left|\{\gamma \in \Gamma : d_X(x,\gamma x) \leqslant r\}\right| \leqslant Ce^{Cr}. \tag{4-26}$$

We claim that there exist $c > 0$, $C > 0$ and $N \in \mathbb{N}$ such that, for $t > 0$ and $x, x' \in X$, we have

$$|p_t^{X,\tau}(x,x')| \leqslant \frac{C}{t^N}\exp\left(-c\frac{d_X^2(x,x')}{t} + Ct\right). \tag{4-27}$$

Indeed, if $\tau = \mathbf{1}$, then $p_t^{X,\mathbf{1}}(x,x')$ is the heat kernel for the Laplace–Beltrami operator. In this case, (4-27) is a consequence of the Li–Yau estimate [1986, Corollary 3.1] and of the fact that X is a symmetric space.

For general τ, using the Itô formula as in [Bismut and Zhang 1992, Equation (12.30)], we can show that there is $C > 0$ such that

$$|p_t^{X,\tau}(x,x')| \leqslant Ce^{Ct} p_t^{X,\mathbf{1}}(x,x'), \tag{4-28}$$

from which we get (4-27).[2]

Note that there exists $c_0 > 0$ such that, for all $\gamma \in \Gamma - \{1\}$ and $x \in X$,

$$d_X(x, \gamma x) \geqslant c_0. \tag{4-29}$$

By (4-27) and (4-29), there exist $c_1 > 0,\ c_2 > 0$ and $C > 0$ such that, for $t > 0,\ x \in X$ and $\gamma \in \Gamma - \{1\}$, we have

$$|p_t^{X,\tau}(x,\gamma x)| \leqslant C \exp\left(-\frac{c_1}{t} - c_2 \frac{d_X^2(x,\gamma x)}{t} + Ct\right). \tag{4-30}$$

By (4-26) and (4-30), for $t > 0$ and $x \in X$, we have

$$\begin{aligned}
\sum_{\gamma \in \Gamma - \{1\}} |p_t^{X,\tau}(x,\gamma x)| &\leqslant C \sum_{\gamma\in\Gamma} \exp\left(-\frac{c_1}{t} - c_2 \frac{d_X^2(x,\gamma x)}{t} + Ct\right) \\
&= c_2 C \exp\left(-\frac{c_1}{t} + Ct\right) \sum_{\gamma \in \Gamma} \int_{d_X^2(x,\gamma x)/t}^{\infty} \exp(-c_2 r)\, dr \\
&= c_2 C \exp\left(-\frac{c_1}{t} + Ct\right) \int_0^{\infty} \left|\{\gamma \in \Gamma : d_X(x,\gamma x) \leqslant \sqrt{rt}\}\right| \exp(-c_2 r)\, dr \\
&\leqslant C' \exp\left(-\frac{c_1}{t} + Ct\right) \int_0^\infty \exp(-c_2 r + C\sqrt{rt})\, dr.
\end{aligned} \tag{4-31}$$

From (4-31), we get (4-25). ☐

For $\gamma \in \Gamma$, set

$$\Gamma(\gamma) = Z(\gamma) \cap \Gamma. \tag{4-32}$$

Let $[\gamma]$ be the conjugacy class of γ in Γ. Let $[\Gamma]$ be the set of all the conjugacy classes of Γ.

The following proposition is [Selberg 1960, Lemma 2]. We include a proof for the sake of completeness.

Proposition 4.9. *If $\gamma \in \Gamma$, then $\Gamma(\gamma)$ is cocompact in $Z(\gamma)$.*

Proof. Since Γ is discrete, $[\gamma]$ is closed in G. The inverse image of $[\gamma]$ by the continuous map $g \in G \to g\gamma g^{-1} \in G$ is $\Gamma \cdot Z(\gamma)$. Then $\Gamma \cdot Z(\gamma)$ is closed in G. Since $\Gamma\backslash G$ is compact, the closed subset $\Gamma\backslash\Gamma \cdot Z(\gamma) \subset \Gamma\backslash G$ is then compact.

The group $Z(\gamma)$ acts transitively on the right on $\Gamma\backslash\Gamma \cdot Z(\gamma)$. The stabilizer at $[1] \in \Gamma\backslash\Gamma\cdot Z(\gamma)$ is $\Gamma(\gamma)$. Hence $\Gamma(\gamma)\backslash Z(\gamma) \simeq \Gamma\backslash\Gamma\cdot Z(\gamma)$ is compact. ☐

Let $\operatorname{vol}(\Gamma(\gamma)\backslash X(\gamma))$ be the volume of $\Gamma(\gamma)\backslash X(\gamma)$ with respect to the volume form induced by $dv_{X(\gamma)}$. Clearly, $\operatorname{vol}(\Gamma(\gamma)\backslash X(\gamma))$ depends only on the conjugacy class $[\gamma] \in [\Gamma]$.

[2] See [Ma and Marinescu 2015, Theorem 4] for another proof of (4-27) using finite propagation speed of solutions of hyperbolic equations.

By the property of heat kernels on compact manifolds, the operator $\exp(-tC^{\mathfrak{g},Z,\tau,\rho}/2)$ is trace class. Its trace is given by the Selberg trace formula:

Theorem 4.10. *There exist $c>0$, $C>0$ such that, for $t>0$, we have*

$$\sum_{[\gamma]\in[\Gamma]-\{1\}} \operatorname{vol}(\Gamma(\gamma)\backslash X(\gamma)) \left|\operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]\right| \leqslant C\exp\Big(-\frac{c}{t}+Ct\Big). \tag{4-33}$$

For $t>0$, the following identity holds:

$$\operatorname{Tr}[\exp(-tC^{\mathfrak{g},Z,\tau,\rho}/2)] = \sum_{[\gamma]\in[\Gamma]} \operatorname{vol}(\Gamma(\gamma)\backslash X(\gamma)) \operatorname{Tr}[\rho(\gamma)] \operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]. \tag{4-34}$$

Proof. Let $F\subset X$ be a fundamental domain of Z in X. By [Bismut 2011, Equations (4.8.11), (4.8.15)], we have

$$\sum_{\gamma'\in[\gamma]} \int_{x\in F} \operatorname{Tr}^{E_\tau}[p_t^{X,\tau}(x,\gamma' x)]\,dx = \operatorname{vol}(\Gamma(\gamma)\backslash X(\gamma)) \operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]. \tag{4-35}$$

By (4-25) and (4-35), we get (4-33). The proof of (4-34) is well known; see, for example, [Bismut 2011, Section 4.8]. □

4E. *A formula for* $\operatorname{Tr}_s^{[\gamma]}[N^{\Lambda^\bullet(T^*X)}\exp(-tC^{\mathfrak{g},X}/2)]$. Let $\gamma=e^a k^{-1}\in G$ be semisimple such that (3-9) holds. Let $\mathfrak{t}(\gamma)\subset\mathfrak{k}(\gamma)$ be a Cartan subalgebra of $\mathfrak{k}(\gamma)$. Set

$$\mathfrak{b}(\gamma)=\{Y\in\mathfrak{p}:\operatorname{Ad}(k)Y=Y,\ [Y,\mathfrak{t}(\gamma)]=0\}. \tag{4-36}$$

Then,

$$a\in\mathfrak{b}(\gamma). \tag{4-37}$$

By definition, $\dim\mathfrak{p}-\dim\mathfrak{b}(\gamma)$ is even.

Since k centralizes $\mathfrak{t}(\gamma)$, by [Knapp 1986, Theorem 4.21], there is $k'\in K$ such that

$$k'\mathfrak{t}(\gamma)k'^{-1}\subset\mathfrak{t},\quad k'kk'^{-1}\in T. \tag{4-38}$$

Up to a conjugation on γ, we can assume directly that $\gamma=e^a k^{-1}$ with

$$\mathfrak{t}(\gamma)\subset\mathfrak{t},\quad k\in T. \tag{4-39}$$

By (3-18), (4-36), and (4-39), we have

$$\mathfrak{b}\subset\mathfrak{b}(\gamma). \tag{4-40}$$

Proposition 4.11. *A semisimple element $\gamma\in G$ can be conjugated into H if and only if*

$$\dim\mathfrak{b}=\dim\mathfrak{b}(\gamma). \tag{4-41}$$

Proof. If $\gamma\in H$, then $\mathfrak{t}(\gamma)=\mathfrak{t}$. By (4-36), we get $\mathfrak{b}=\mathfrak{b}(\gamma)$, which implies (4-41).

Recall that $\mathfrak{h}(\gamma)\subset\mathfrak{z}(\gamma)$ is defined as in (3-20), when G is replaced by $Z^0(\gamma)$ and $\mathfrak{t}$ is replaced by $\mathfrak{t}(\gamma)$. It is a θ-invariant Cartan subalgebra of both $\mathfrak{g}$ and $\mathfrak{z}(\gamma)$. Let $\mathfrak{h}(\gamma)=\mathfrak{h}(\gamma)_{\mathfrak{p}}\oplus\mathfrak{h}(\gamma)_{\mathfrak{k}}$ be the Cartan decomposition. Then,

$$\mathfrak{h}(\gamma)_{\mathfrak{p}}=\{Y\in\mathfrak{p}(\gamma):[Y,\mathfrak{t}(\gamma)]=0\}=\mathfrak{b}(\gamma)\cap\mathfrak{p}(\gamma),\quad \mathfrak{h}(\gamma)_{\mathfrak{k}}=\mathfrak{t}(\gamma). \tag{4-42}$$

From (3-26) and (4-42), we get

$$\dim\mathfrak{b}\leqslant\dim\mathfrak{h}(\gamma)_{\mathfrak{p}}\leqslant\dim\mathfrak{b}(\gamma). \tag{4-43}$$

By (4-43), if $\dim\mathfrak{b}=\dim\mathfrak{b}(\gamma)$, then $\dim\mathfrak{b}=\dim\mathfrak{h}(\gamma)_{\mathfrak{p}}$. By Proposition 3.3, γ can be conjugated into H. □

The following theorem extends [Bismut 2011, Theorem 7.9.1].

Theorem 4.12. *Let $\gamma\in G$ be semisimple such that* $\dim\mathfrak{b}(\gamma)\geqslant 2$. *For $Y\in\mathfrak{k}(\gamma)$, we have*

$$\mathrm{Tr_s}^{\Lambda^\bullet(\mathfrak{p}^*)}[N^{\Lambda^\bullet(\mathfrak{p}^*)}\,\mathrm{Ad}(k^{-1})\exp(-i\,\mathrm{ad}(Y))]=0. \tag{4-44}$$

In particular, for $t>0$, we have

$$\mathrm{Tr_s}^{[\gamma]}[N^{\Lambda^\bullet(T^*X)}\exp(-tC^{\mathfrak{g},X}/2)]=0. \tag{4-45}$$

Proof. Since the left-hand side of (4-44) is $\mathrm{Ad}(K^0(\gamma))$-invariant, it is enough to show (4-44) for $Y\in\mathfrak{t}(\gamma)$. If $Y\in\mathfrak{t}(\gamma)$, by [Bismut 2011, Equation (7.9.1)], we have

$$\mathrm{Tr_s}^{\Lambda^\bullet(\mathfrak{p}^*)}[N^{\Lambda^\bullet(\mathfrak{p}^*)}\,\mathrm{Ad}(k^{-1})\exp(-i\,\mathrm{ad}(Y))]=\frac{\partial}{\partial b}\bigg|_{b=0}\det\big(1-e^b\,\mathrm{Ad}(k)\exp(i\,\mathrm{ad}(Y))\big)\big|_{\mathfrak{p}}. \tag{4-46}$$

Since $\dim\mathfrak{b}(\gamma)\geqslant 2$, by (4-46), we get (4-44) for $Y\in\mathfrak{t}(\gamma)$.

By (4-22) and (4-44), we get (4-45). □

In this way, [Bismut 2011, Theorem 7.9.3] recovered [Moscovici and Stanton 1991, Corollary 2.2].

Corollary 4.13. *Let F be a unitarily flat vector bundle on Z. Assume that* $\dim Z$ *is odd and* $\delta(G)\neq 1$. *Then for any $t>0$, we have*

$$\mathrm{Tr_s}[N^{\Lambda^\bullet(T^*Z)}\exp(-t\square^Z/2)]=0. \tag{4-47}$$

In particular,

$$T(F)=1. \tag{4-48}$$

Proof. Since $\dim Z$ is odd, $\delta(G)$ is odd. Since $\delta(G)\neq 1$, we have $\delta(G)\geqslant 3$. By (4-40), $\dim\mathfrak{b}(\gamma)\geqslant\delta(G)\geqslant 3$, so (4-47) is a consequence of (4-24), (4-34) and (4-45). □

Suppose $\delta(G)=1$. Up to sign, we fix an element $a_1\in\mathfrak{b}$ such that $B(a_1,a_1)=1$. As in Section 3B, set

$$M=Z^{a_1,\perp,0}(e^{a_1}),\quad K_M=K^0(e^{a_1}), \tag{4-49}$$

and

$$\mathfrak{m}=\mathfrak{z}^{a_1,\perp}(e^{a_1}),\quad \mathfrak{p}_{\mathfrak{m}}=\mathfrak{p}^{a_1,\perp}(e^{a_1}),\quad \mathfrak{k}_{\mathfrak{m}}=\mathfrak{k}(e^{a_1}). \tag{4-50}$$

As in Section 3B, M is a connected reductive group with Lie algebra $\mathfrak{m}$, with maximal compact subgroup K_M, and with Cartan decomposition $\mathfrak{m} = \mathfrak{p}_\mathfrak{m} \oplus \mathfrak{k}_\mathfrak{m}$. Let

$$X_M = M/K_M \tag{4-51}$$

be the corresponding symmetric space. By definition, $T \subset M$ is a compact Cartan subgroup. Therefore $\delta(M) = 0$, and $\dim \mathfrak{p}_\mathfrak{m}$ is even.

Assume that $\delta(G) = 1$ and that G has noncompact center, so that $\dim \mathfrak{z}_\mathfrak{p} \geqslant 1$. By (3-19), we find that $a_1 \in \mathfrak{z}_\mathfrak{p}$, so that $Z^0(a_1) = G$. By (3-14) and (4-10), we have

$$G = \mathbb{R} \times M, \quad K = K_M, \quad X = \mathbb{R} \times X_M. \tag{4-52}$$

Let $\gamma \in G$ be a semisimple element such that $\dim \mathfrak{b}(\gamma) = 1$. By Proposition 4.11, we may assume that $\gamma = e^a k^{-1}$ with $a \in \mathfrak{b}$ and $k \in T$.

Proposition 4.14. *We have*

$$\mathrm{Tr_s}^{[1]}[N^{\Lambda^\bullet(T^*X)} \exp(-tC^{\mathfrak{g},X}/2)] = -\frac{1}{\sqrt{2\pi t}}[e(TX_M, \nabla^{TX_M})]^{\max}. \tag{4-53}$$

If $\gamma = e^a k^{-1}$ with $a \in \mathfrak{b}$, $a \neq 0$, and $k \in T$, then

$$\mathrm{Tr}^{[\gamma]}[N^{\Lambda^\bullet(T^*X)} \exp(-tC^{\mathfrak{g},X}/2)] = -\frac{1}{\sqrt{2\pi t}} e^{-\frac{|a|^2}{2t}} [e(TX^{a,\perp}(\gamma), \nabla^{TX^{a,\perp}(\gamma)})]^{\max}. \tag{4-54}$$

Proof. By (4-52), for $\gamma = e^a k^{-1}$ with $a \in \mathfrak{b}$ and $k \in T$, we have

$$\mathrm{Tr}_s^{[\gamma]}[N^{\Lambda^\bullet(T^*X)} \exp(-tC^{\mathfrak{g},X}/2)] = -\mathrm{Tr}^{[e^a]}[\exp(t\Delta^{\mathbb{R}}/2)]\, \mathrm{Tr}_s^{[k^{-1}]}[\exp(-tC^{\mathfrak{m},X_M}/2)], \tag{4-55}$$

where $\Delta^{\mathbb{R}}$ is the Laplace–Beltrami operator acting on $C^\infty(\mathbb{R})$.

Clearly,

$$\mathrm{Tr}^{[e^a]}[\exp(t\Delta^{\mathbb{R}}/2)] = \frac{1}{\sqrt{2\pi t}} e^{-\frac{|a|^2}{2t}}. \tag{4-56}$$

By [Bismut 2011, Theorem 7.8.13], we have

$$\mathrm{Tr}_s^{[1]}[\exp(-tC^{\mathfrak{m},X_M}/2)] = [e(TX_M, \nabla^{TX_M})]^{\max} \tag{4-57}$$

and

$$\mathrm{Tr}_s^{[k^{-1}]}[\exp(-tC^{\mathfrak{m},X_M}/2)] = [e(TX^{a,\perp}(\gamma), \nabla^{TX^{a,\perp}(\gamma)})]^{\max}. \tag{4-58}$$

By (4-55)–(4-58), we get (4-53) and (4-54). □

5. The solution to Fried conjecture

We use the notation in Sections 3 and 4. Also, we assume that $\dim \mathfrak{p}$ is odd. The purpose of this section is to introduce the Ruelle dynamical zeta function on Z and to state our main result, which contains the solution of the Fried conjecture in the case of locally symmetric spaces.

This section is organized as follows. In Section 5A, we describe the closed geodesics on Z.

In Section 5B, we define the dynamical zeta function and state Theorem 5.5, which is the main result of the article.

Finally, in Section 5C, we establish Theorem 5.5 when G has noncompact center and $\delta(G) = 1$.

5A. ***The space of closed geodesics.*** By [Duistermaat et al. 1979, Proposition 5.15], the set of nontrivial closed geodesics on Z consists of a disjoint union of smooth connected closed submanifolds

$$\coprod_{[\gamma]\in[\Gamma]-[1]} B_{[\gamma]}. \tag{5-1}$$

Moreover, $B_{[\gamma]}$ is diffeomorphic to $\Gamma(\gamma)\backslash X(\gamma)$. All the elements of $B_{[\gamma]}$ have the same length $|a| > 0$ if γ can be conjugated to $e^a k^{-1}$ as in (3-9). Also, the geodesic flow induces a canonical locally free action of $\mathbb{S}^1$ on $B_{[\gamma]}$, so that $\mathbb{S}^1\backslash B_{[\gamma]}$ is a closed orbifold. The $\mathbb{S}^1$-action is not necessarily effective. Let

$$m_{[\gamma]} = \left|\ker(\mathbb{S}^1 \to \mathrm{Diff}(B_{[\gamma]}))\right| \in \mathbb{N}^* \tag{5-2}$$

be the generic multiplicity.

Following [Satake 1957], if S is a closed Riemannian orbifold with Levi-Civita connection ∇^{TS}, then $e(TS, \nabla^{TS}) \in \Omega^{\dim S}(S, o(TS))$ is still well defined, and the Euler characteristic $\chi_{\mathrm{orb}}(S) \in \mathbb{Q}$ is given by

$$\chi_{\mathrm{orb}}(S) = \int_S e(TS, \nabla^{TS}). \tag{5-3}$$

Proposition 5.1. *For $\gamma \in \Gamma - \{1\}$, the following identity holds*:

$$\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}} = \frac{\mathrm{vol}(\Gamma(\gamma)\backslash X(\gamma))}{|a|}\left[e(TX^{a,\perp}(\gamma), \nabla^{TX^{a,\perp}(\gamma)})\right]^{\max}. \tag{5-4}$$

Proof. Take $\gamma \in \Gamma - \{1\}$. We can assume that $\gamma = e^a k^{-1}$ as in (3-9) with $a \neq 0$. By (3-10) and (4-32), for $t \in \mathbb{R}$, we know e^{ta} commutes with elements of $\Gamma(\gamma)$. Thus, e^{ta} acts on the left on $\Gamma(\gamma)\backslash X(\gamma)$. Since $e^a = \gamma k,\ \gamma \in \Gamma(\gamma),\ k \in K(\gamma)$ and k commutes with elements of $Z(\gamma)$, we see that e^a acts as identity on $\Gamma(\gamma)\backslash X(\gamma)$. This induces an $\mathbb{R}/\mathbb{Z} \simeq \mathbb{S}^1$ action on $\Gamma(\gamma)\backslash X(\gamma)$ which coincides with the $\mathbb{S}^1$-action on $B_{[\gamma]}$. Therefore,

$$\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]}) = \mathrm{vol}(\mathbb{S}^1\backslash B_{[\gamma]})\left[e(TX^{a,\perp}(\gamma), \nabla^{TX^{a,\perp}(\gamma)})\right]^{\max} \tag{5-5}$$

and

$$\frac{\mathrm{vol}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}} = \frac{\mathrm{vol}(\Gamma(\gamma)\backslash X(\gamma))}{|a|}. \tag{5-6}$$

By (5-5) and (5-6), we get (5-4). □

Corollary 5.2. *Let $\gamma \in \Gamma - \{1\}$. If* $\dim \mathfrak{b}(\gamma) \geqslant 2$, *then*

$$\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]}) = 0. \tag{5-7}$$

Proof. By Propositions 4.1 and 5.1, it is enough to show that

$$\delta(Z^{a,\perp,0}(\gamma)) \geqslant 1. \tag{5-8}$$

By (3-14) and (3-26), we have

$$\delta(Z^{a,\perp,0}(\gamma)) = \delta(Z^0(\gamma)) - 1 \geqslant \delta(G) - 1. \tag{5-9}$$

Recall $\dim \mathfrak{p}$ is odd, therefore $\delta(G)$ is odd. If $\delta(G) \geqslant 3$, by (5-9), we get (5-8). If $\delta(G) = 1$, then $\dim \mathfrak{b}(\gamma) \geqslant 2 > \delta(G)$. By Propositions 3.3 and 4.11, the inequality in (5-9) is strict, which implies (5-8). □

Remark 5.3. By Theorem 4.12 and Corollary 5.2, we know both $\mathrm{Tr_s}^{[\gamma]}[N^{\Lambda^\bullet(T^*X)} \exp(-tC^{\mathfrak{g},X}/2)]$ and $\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})$ vanish when $\dim \mathfrak{b}(\gamma) \geqslant 2$.

5B. ***Statement of the main result.*** Recall that $\rho : \Gamma \to \mathrm{U}(r)$ is a unitary representation of Γ and that (F, ∇^F, g^F) is the unitarily flat vector bundle on Z associated with ρ.

Definition 5.4. The Ruelle dynamical zeta function $R_\rho(\sigma)$ is said to be well defined, if the following properties hold:

(1) For $\sigma \in \mathbb{C}$, $\mathrm{Re}(\sigma) \gg 1$, the sum

$$\Xi_\rho(\sigma) = \sum_{[\gamma]\in[\Gamma]-\{1\}} \frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}} \mathrm{Tr}[\rho(\gamma)] e^{-\sigma|a|} \tag{5-10}$$

converges to a holomorphic function.

(2) The function $R_\rho(\sigma) = \exp(\Xi_\rho(\sigma))$ has a meromorphic extension to $\sigma \in \mathbb{C}$.

If $\delta(G) \neq 1$, by Corollary 5.2,

$$R_\rho(\sigma) \equiv 1. \tag{5-11}$$

The main result of this article is the solution of the Fried conjecture. We restate Theorem 1.1 as follows.

Theorem 5.5. *The dynamical zeta function $R_\rho(\sigma)$ is well defined. There exist explicit constants $C_\rho \in \mathbb{R}^*$ and $r_\rho \in \mathbb{Z}$, see* (7-75), *such that, when $\sigma \to 0$ we have*

$$R_\rho(\sigma) = C_\rho T(F)^2 \sigma^{r_\rho} + \mathcal{O}(\sigma^{r_\rho+1}). \tag{5-12}$$

If $H^\bullet(Z, F) = 0$, then

$$C_\rho = 1, \quad r_\rho = 0, \tag{5-13}$$

so that

$$R_\rho(0) = T(F)^2. \tag{5-14}$$

Proof. When $\delta(G) \neq 1$, Theorem 5.5 is a consequence of (4-48) and (5-11). When $\delta(G) = 1$ and when G has noncompact center, we will show Theorem 5.5 in Section 5C. When $\delta(G) = 1$ and when G has compact center, we will show that $R_\rho(\sigma)$ is well defined such that (5-12) holds in Section 7, and we will show (5-13) in Section 8. □

5C. *Proof of Theorem 5.5 when G has noncompact center and $\delta(G) = 1$.* We assume that $\delta(G) = 1$ and that G has noncompact center. Let us show the following refined version of Theorem 5.5.

Theorem 5.6. *There is $\sigma_0 > 0$ such that*

$$\sum_{[\gamma]\in[\Gamma]-\{1\}} \frac{|\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})|}{m_{[\gamma]}} e^{-\sigma_0|a|} < \infty. \tag{5-15}$$

The dynamical zeta function $R_\rho(\sigma)$ extends meromorphically to $\sigma \in \mathbb{C}$ such that

$$R_\rho(\sigma) = \exp\bigl(r\,\mathrm{vol}(Z)[e(TX_M, \nabla^{TX_M})]^{\max}\sigma\bigr)T(\sigma^2). \tag{5-16}$$

If $\chi'(Z,F) = 0$, then $R_\rho(\sigma)$ is holomorphic at $\sigma = 0$ and

$$R_\rho(0) = T(F)^2. \tag{5-17}$$

Proof. Following (2-5), for $(s,\sigma) \in \mathbb{C}\times\mathbb{R}$ such that $\mathrm{Re}(s) > m/2$ and $\sigma > 0$, put

$$\begin{aligned}\theta_\rho(s,\sigma) &= -\operatorname{Tr}[N^{\Lambda^\bullet(T^*Z)}(C^{\mathfrak{g},Z,\rho}+\sigma)^{-s}]\\ &= -\frac{1}{\Gamma(s)}\int_0^\infty \operatorname{Tr_s}\bigl[N^{\Lambda^\bullet(T^*Z)}\exp(-t(C^{\mathfrak{g},Z,\rho}+\sigma))\bigr]t^{s-1}\,dt.\end{aligned} \tag{5-18}$$

Let us show that there is $\sigma_0 > 0$ such that (5-15) holds true and that for $\sigma > \sigma_0$, we have

$$\Xi_\rho(\sigma) = \frac{\partial}{\partial s}\theta_\rho(0,\sigma^2) + r\,\mathrm{vol}(Z)[e(TX_M,\nabla^{TX_M})]^{\max}\sigma. \tag{5-19}$$

By (4-53), for $(s,\sigma)\in\mathbb{C}\times\mathbb{R}$ such that $\mathrm{Re}(s) > \frac{1}{2}$ and $\sigma > 0$, the function

$$\theta_{\rho,1}(s,\sigma) = -\frac{r\,\mathrm{vol}(Z)}{\Gamma(s)}\int_0^\infty \operatorname{Tr_s}^{[1]}\bigl[N^{\Lambda^\bullet(T^*X)}\exp(-t(C^{\mathfrak{g},X}+\sigma))\bigr]t^{s-1}\,dt \tag{5-20}$$

is well defined so that

$$\theta_{\rho,1}(s,\sigma) = \frac{r\,\mathrm{vol}(Z)}{2\sqrt{\pi}}[e(TX_M,\nabla^{TX_M})]^{\max}\frac{\Gamma\bigl(s-\frac{1}{2}\bigr)}{\Gamma(s)}\sigma^{\frac{1}{2}-s}. \tag{5-21}$$

Therefore, for $\sigma > 0$ fixed, the function $s \to \theta_{\rho,1}(s,\sigma)$ has a meromorphic extension to $s\in\mathbb{C}$ which is holomorphic at $s = 0$ so that

$$\frac{\partial}{\partial s}\theta_{\rho,1}(0,\sigma) = -r\,\mathrm{vol}(Z)[e(TX_M,\nabla^{TX_M})]^{\max}\sigma^{\frac{1}{2}}. \tag{5-22}$$

For $(s,\sigma)\in\mathbb{C}\times\mathbb{R}$ such that $\mathrm{Re}(s) > m/2$ and $\sigma > 0$, set

$$\theta_{\rho,2}(s,\sigma) = \theta_\rho(s,\sigma) - \theta_{\rho,1}(s,\sigma). \tag{5-23}$$

By (4-45), (4-54), (5-4), and (5-7), for $[\gamma]\in[\Gamma]-\{1\}$, we have

$$\mathrm{vol}(\Gamma(\gamma)\backslash X(\gamma))\operatorname{Tr_s}^{[\gamma]}[N^{\Lambda^\bullet(T^*X)}\exp(-tC^{\mathfrak{g},X})] = -\frac{1}{2\sqrt{\pi t}}\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}|a|\exp\Bigl(-\frac{|a|^2}{4t}\Bigr). \tag{5-24}$$

By (4-33) and (5-24), there exist $C_1 > 0$, $C_2 > 0$, and $C_3 > 0$ such that, for $t > 0$, we have

$$\sum_{[\gamma]\in[\Gamma]-\{1\}} \frac{|\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})|}{m_{[\gamma]}}|a|\exp\left(-\frac{|a|^2}{4t}\right) \leqslant C_1\exp\left(-\frac{C_2}{t}+C_3t\right). \tag{5-25}$$

Take $\sigma_0 = \sqrt{2C_3}$. Since ρ is unitary, by (4-34), (5-18), (5-23), and (5-25), for $(s,\sigma)\in\mathbb{C}\times\mathbb{R}$ such that $\mathrm{Re}(s) > m/2$ and $\sigma \geqslant \sigma_0$, we have

$$\theta_{\rho,2}(s,\sigma^2) = \frac{1}{2\sqrt{\pi}\Gamma(s)}\int_0^\infty \sum_{[\gamma]\in[\Gamma]-\{1\}} \mathrm{Tr}[\rho(\gamma)]\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}|a|\exp\left(-\frac{|a|^2}{4t}-\sigma^2t\right)t^{s-\frac{3}{2}}\,dt. \tag{5-26}$$

Moreover, for $\sigma \geqslant \sigma_0$ fixed, the function $s \to \theta_{\rho,2}(s,\sigma^2)$ extends holomorphically to $\mathbb{C}$, so that

$$\frac{\partial}{\partial s}\theta_{\rho,2}(0,\sigma^2) = \frac{1}{2\sqrt{\pi}}\int_0^\infty \sum_{[\gamma]\in[\Gamma]-\{1\}} \mathrm{Tr}[\rho(\gamma)]\frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}}|a|\exp\left(-\frac{|a|^2}{4t}-\sigma^2t\right)\frac{dt}{t^{\frac{3}{2}}}. \tag{5-27}$$

Using the formula[3] that for $B_1 > 0$, $B_2 \geqslant 0$,

$$\int_0^\infty \exp\left(-\frac{B_1}{t}-B_2t\right)\frac{dt}{t^{\frac{3}{2}}} = \sqrt{\frac{\pi}{B_1}}\exp(-2\sqrt{B_1B_2}), \tag{5-28}$$

by (5-25), (5-27), and by Fubini's theorem, we get (5-15). Also, for $\sigma \geqslant \sigma_0$, we have

$$\frac{\partial}{\partial s}\theta_{\rho,2}(0,\sigma^2) = \Xi_\rho(\sigma). \tag{5-29}$$

By (5-22), (5-23), and (5-29), we get (5-19). By taking the exponentials, we get (5-16) for $\sigma \geqslant \sigma_0$. Since the right-hand side of (5-16) is meromorphic on $\sigma\in\mathbb{C}$, we know R_ρ has a meromorphic extension to $\mathbb{C}$. By (2-15) and (5-16), we get (5-17). □

6. Reductive groups G with compact center and $\delta(G) = 1$

In this section, we assume that $\delta(G) = 1$ and that G has compact center. The purpose of this section is to introduce some geometric objects associated with G. Their properties are proved by algebraic arguments based on the classification of real simple Lie algebras $\mathfrak{g}$ with $\delta(\mathfrak{g}) = 1$. The results of this section will be used in Section 7, in order to evaluate certain orbital integrals.

This section is organized as follows. In Section 6A, we introduce a splitting $\mathfrak{g} = \mathfrak{b}\oplus\mathfrak{m}\oplus\mathfrak{n}\oplus\bar{\mathfrak{n}}$, associated with the action of $\mathfrak{b}$ on $\mathfrak{g}$.

In Section 6B, we construct a natural compact Hermitian symmetric space $Y_{\mathfrak{b}}$, which will be used in the calculation of orbital integrals in Section 7A.

In Section 6C, we state one key result, which says that the action of K_M on $\mathfrak{n}$ lifts to K. The purpose of the following subsections is to prove this result.

[3] We give a proof of (5-28) when $B_1 = B_2 = 1$. Indeed, we have $\int_0^\infty \exp\left(-\frac{1}{t}-t\right)\frac{dt}{t^{3/2}} = \frac{1}{2}\int_0^\infty \exp\left(-\frac{1}{t}-t\right)\left(\frac{1}{t^{3/2}}+\frac{1}{t^{1/2}}\right)dt$. Using the change of variables $u = t^{1/2}-t^{-1/2}$, we get (5-28).

In Section 6D, we state a classification result of real simple Lie algebras $\mathfrak{g}$ with $\delta(\mathfrak{g}) = 1$, which asserts that they just contain $\mathfrak{sl}_3(\mathbb{R})$ and $\mathfrak{so}(p,q)$ with $pq > 1$ odd. This result has already been used by Moscovici and Stanton [1991].

In Sections 6E and 6F, we study the Lie groups $\mathrm{SL}_3(\mathbb{R})$ and $\mathrm{SO}^0(p,q)$ with $pq > 1$ odd, and the structure of the associated Lie groups M, K_M.

In Section 6G, we study the connected component G_* of the identity of the isometry group of $X = G/K$. We show that G_* has a factor $\mathrm{SL}_3(\mathbb{R})$ or $\mathrm{SO}^0(p,q)$ with $pq > 1$ odd.

Finally, in Sections 6H–6L, we show several unproven results stated in Sections 6A–6C. Most of the results are shown case by case for the groups $\mathrm{SL}_3(\mathbb{R})$ and $\mathrm{SO}^0(p,q)$ with $pq > 1$ odd. We prove the corresponding results for general G using a natural morphism $i_G : G \to G_*$.

6A. ***A splitting of*** $\mathfrak{g}$**.** We use the notation in (4-49)–(4-51). Let $Z(\mathfrak{b}) \subset G$ be the stabilizer of $\mathfrak{b}$ in G, and let $\mathfrak{z}(\mathfrak{b}) \subset \mathfrak{g}$ be its Lie algebra.

We define $\mathfrak{p}(\mathfrak{b})$, $\mathfrak{k}(\mathfrak{b})$, $\mathfrak{p}^\perp(\mathfrak{b})$, $\mathfrak{k}^\perp(\mathfrak{b})$, $\mathfrak{z}^\perp(\mathfrak{b})$ in an obvious way as in Section 3B. By (4-50), we have

$$\mathfrak{p}(\mathfrak{b}) = \mathfrak{b} \oplus \mathfrak{p}_\mathfrak{m}, \quad \mathfrak{k}(\mathfrak{b}) = \mathfrak{k}_\mathfrak{m}. \tag{6-1}$$

Also,

$$\mathfrak{p} = \mathfrak{b} \oplus \mathfrak{p}_\mathfrak{m} \oplus \mathfrak{p}^\perp(\mathfrak{b}), \quad \mathfrak{k} = \mathfrak{k}_\mathfrak{m} \oplus \mathfrak{k}^\perp(\mathfrak{b}). \tag{6-2}$$

Let $Z^0(\mathfrak{b})$ be the connected component of the identity in $Z(\mathfrak{b})$. By (3-14), we have

$$Z^0(\mathfrak{b}) = \mathbb{R} \times M. \tag{6-3}$$

The group K_M acts trivially on $\mathfrak{b}$. It also acts on $\mathfrak{p}_\mathfrak{m}$, $\mathfrak{p}^\perp(\mathfrak{b})$, $\mathfrak{k}_\mathfrak{m}$ and $\mathfrak{k}^\perp(\mathfrak{b})$, and preserves the splittings (6-2).

Recall that we have fixed $a_1 \in \mathfrak{b}$ such that $B(a_1, a_1) = 1$. The choice of a_1 fixes an orientation of $\mathfrak{b}$. Let $\mathfrak{n} \subset \mathfrak{z}^\perp(\mathfrak{b})$ be the direct sum of the eigenspaces of $\mathrm{ad}(a_1)$ with the positive eigenvalues. Set $\bar{\mathfrak{n}} = \theta\mathfrak{n}$. Then $\bar{\mathfrak{n}}$ is the direct sum of the eigenspaces with negative eigenvalues, and

$$\mathfrak{z}^\perp(\mathfrak{b}) = \mathfrak{n} \oplus \bar{\mathfrak{n}}. \tag{6-4}$$

Clearly, $Z^0(\mathfrak{b})$ acts on $\mathfrak{n}$ and $\bar{\mathfrak{n}}$ by adjoint action. Since K_M is fixed by θ, we have isomorphisms of representations of K_M

$$X \in \mathfrak{n} \to X - \theta X \in \mathfrak{p}^\perp(\mathfrak{b}), \quad X \in \mathfrak{n} \to X + \theta X \in \mathfrak{k}^\perp(\mathfrak{b}). \tag{6-5}$$

In the sequel, if $f \in \mathfrak{n}$, we define $\bar{f} = \theta f \in \bar{\mathfrak{n}}$.

By (6-2) and (6-5), we have $\dim \mathfrak{n} = \dim \mathfrak{p} - \dim \mathfrak{p}_\mathfrak{m} - 1$. Since $\dim \mathfrak{p}$ is odd and since $\dim \mathfrak{p}_\mathfrak{m}$ is even, $\dim \mathfrak{n}$ is even. Set

$$l = \tfrac{1}{2} \dim \mathfrak{n}. \tag{6-6}$$

Note that since G has compact center, we have $\mathfrak{b} \not\subset \mathfrak{z}_\mathfrak{g}$. Therefore, $\mathfrak{z}^\perp(\mathfrak{b}) \neq 0$ and $l > 0$.

Remark 6.1. Let $\mathfrak{q} \subset \mathfrak{g}$ be the direct sum of the eigenspaces of $\mathrm{ad}(a_1)$ with nonnegative eigenvalues. Then $\mathfrak{q}$ is a proper parabolic subalgebra of $\mathfrak{g}$, with Langlands decomposition $\mathfrak{q} = \mathfrak{m} \oplus \mathfrak{b} \oplus \mathfrak{n}$ [Knapp 2002, Section VII.7]. Let $Q \subset G$ be the corresponding parabolic subgroup of G, and let $Q = M_Q A_Q N_Q$ be the corresponding Langlands decomposition. Then M is the connected component of the identity in M_Q, and $\mathfrak{b}, \mathfrak{n}$ are the Lie algebras of A_Q and N_Q.

Proposition 6.2. *Any element of* $\mathfrak{b}$ *acts on* $\mathfrak{n}$ *and* $\bar{\mathfrak{n}}$ *as a scalar*; *i.e., there exists* $\alpha \in \mathfrak{b}^*$ *such that, for* $a \in \mathfrak{b}$, $f \in \mathfrak{n}$, *we have*

$$[a, f] = \langle \alpha, a \rangle f, \quad [a, \bar{f}] = -\langle \alpha, a \rangle \bar{f}. \tag{6-7}$$

Proof. The proof of Proposition 6.2, based on the classification theory of real simple Lie algebras, will be given in Section 6H. □

Let $a_0 \in \mathfrak{b}$ be such that

$$\langle \alpha, a_0 \rangle = 1. \tag{6-8}$$

Proposition 6.3. *We have*

$$[\mathfrak{n}, \bar{\mathfrak{n}}] \subset \mathfrak{z}(\mathfrak{b}), \quad [\mathfrak{n}, \mathfrak{n}] = [\bar{\mathfrak{n}}, \bar{\mathfrak{n}}] = 0. \tag{6-9}$$

Also,

$$B|_{\mathfrak{n} \times \mathfrak{n} = 0}, \quad B|_{\bar{\mathfrak{n}} \times \bar{\mathfrak{n}}} = 0. \tag{6-10}$$

Proof. By (6-7), $a \in \mathfrak{b}$ acts on $[\mathfrak{n}, \bar{\mathfrak{n}}]$, $[\mathfrak{n}, \mathfrak{n}]$, and $[\bar{\mathfrak{n}}, \bar{\mathfrak{n}}]$ by multiplication by 0, $2\langle \alpha, a \rangle$, and $-2\langle \alpha, a \rangle$. Equation (6-9) follows.

If $f_1, f_2 \in \mathfrak{n}$, by (6-7) and (6-8), we have

$$B(f_1, f_2) = B([a_0, f_1], f_2) = -B(f_1, [a_0, f_2]) = -B(f_1, f_2). \tag{6-11}$$

From (6-11), we get the first equation of (6-10). We obtain the second equation of (6-10) by the same argument. □

Remark 6.4. Clearly, we have

$$[\mathfrak{z}(\mathfrak{b}), \mathfrak{z}(\mathfrak{b})] \subset \mathfrak{z}(\mathfrak{b}). \tag{6-12}$$

Since $\mathfrak{z}(\mathfrak{b})$ preserves B and since $\mathfrak{z}^{\perp}(\mathfrak{b})$ is the orthogonal space to $\mathfrak{z}(\mathfrak{b})$ in $\mathfrak{g}$ with respect to B, we have

$$[\mathfrak{z}(\mathfrak{b}), \mathfrak{z}^{\perp}(\mathfrak{b})] \subset \mathfrak{z}^{\perp}(\mathfrak{b}). \tag{6-13}$$

By (6-4) and (6-9), we get

$$[\mathfrak{z}^{\perp}(\mathfrak{b}), \mathfrak{z}^{\perp}(\mathfrak{b})] \subset \mathfrak{z}(\mathfrak{b}). \tag{6-14}$$

We note the similarity between (3-2) and (6-12)–(6-14). In the sequel, We call such a pair $(\mathfrak{z}, \mathfrak{z}(\mathfrak{b}))$ a symmetric pair.

For $k \in K_M$, let $M(k) \subset M$ be the centralizer of k in M, and let $\mathfrak{m}(k)$ be its Lie algebra. Let $M^0(k)$ be the connected component of the identity in $M(k)$. Let $\mathfrak{p}_{\mathfrak{m}}(k)$ and $\mathfrak{k}_{\mathfrak{m}}(k)$ be the analogues of $\mathfrak{p}(\gamma)$ and $\mathfrak{k}(\gamma)$ in (3-11), so that

$$\mathfrak{m}(k) = \mathfrak{p}_{\mathfrak{m}}(k) \oplus \mathfrak{k}_{\mathfrak{m}}(k). \tag{6-15}$$

Since k is elliptic in M, we know $M^0(k)$ is reductive with maximal compact subgroup $K_M^0(k) = M^0(k) \cap K$ and with Cartan decomposition (6-15). Let

$$X_M(k) = M^0(k)/K_M^0(k) \tag{6-16}$$

be the corresponding symmetric space. Note that $\delta(M^0(k)) = 0$ and $\dim X_M(k)$ is even.

Clearly, if $\gamma = e^a k^{-1} \in H$ with $a \in \mathfrak{b},\ a \neq 0,\ k \in T$, then

$$\mathfrak{p}(\gamma) = \mathfrak{p}_{\mathfrak{m}}(k), \quad \mathfrak{k}(\gamma) = \mathfrak{k}_{\mathfrak{m}}(k), \quad Z^{a,\perp,0}(\gamma) = M^0(k), \quad K^0(\gamma) = K_M^0(k). \tag{6-17}$$

Proposition 6.5. *For $\gamma = e^a k^{-1} \in H$ with $a \in \mathfrak{b},\ a \neq 0,\ k \in T$, we have*

$$\begin{aligned} \left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}} &= \sum_{j=0}^{2l} (-1)^j \operatorname{Tr}^{\Lambda^j(\mathfrak{n}^*)}[\mathrm{Ad}(k^{-1})] e^{(l-j)\langle\alpha,a\rangle} \\ &= \sum_{j=0}^{2l} (-1)^j \operatorname{Tr}^{\Lambda^j(\mathfrak{n}^*)}[\mathrm{Ad}(k^{-1})] e^{(l-j)|\alpha||a|}. \end{aligned} \tag{6-18}$$

Proof. We claim that

$$\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}} = e^{l\langle\alpha,a\rangle} \det(1-\mathrm{Ad}(\gamma))|_{\bar{\mathfrak{n}}}. \tag{6-19}$$

Indeed, since $\dim \bar{\mathfrak{n}}$ is even, the right-hand side of (6-19) is positive. By (6-4), we have

$$\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp} = \det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{n}} \det(1-\mathrm{Ad}(\gamma))|_{\bar{\mathfrak{n}}}. \tag{6-20}$$

Since $\bar{\mathfrak{n}} = \theta\mathfrak{n}$, we have

$$\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{n}} = \det(1-\mathrm{Ad}(\theta\gamma))|_{\bar{\mathfrak{n}}} = \det(\mathrm{Ad}(\theta\gamma))|_{\bar{\mathfrak{n}}} \det(\mathrm{Ad}(\theta\gamma)^{-1}-1)|_{\bar{\mathfrak{n}}}. \tag{6-21}$$

Since $\dim \bar{\mathfrak{n}} = 2l$ is even, and since $(\theta\gamma)^{-1} = e^a k$ and k acts unitarily on $\mathfrak{n}$, by (6-7) and (6-21), we have

$$\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{n}} = e^{2l\langle\alpha,a\rangle} \det(1-\mathrm{Ad}(e^a k))|_{\bar{\mathfrak{n}}} = e^{2l\langle\alpha,a\rangle} \det(1-\mathrm{Ad}(\gamma))|_{\bar{\mathfrak{n}}}. \tag{6-22}$$

By (6-20) and (6-22), we get (6-19).

Classically,

$$\det(1-\mathrm{Ad}(\gamma))|_{\bar{\mathfrak{n}}} = \sum_{j=0}^{2l} (-1)^j \operatorname{Tr}^{\Lambda^j(\bar{\mathfrak{n}})}[\mathrm{Ad}(k^{-1})] e^{-j\langle\alpha,a\rangle}. \tag{6-23}$$

Using the isomorphism of K_M-representations $\mathfrak{n}^* \simeq \bar{\mathfrak{n}}$, by (6-19), (6-23), we get the first equation of (6-18) and the second equation of (6-18) if a is positive in $\mathfrak{b}$. For the case a is negative in $\mathfrak{b}$, it is enough to remark that replacing γ by $\theta\gamma$ does not change the left-hand side of (6-18). □

6B. ***A compact Hermitian symmetric space $Y_{\mathfrak{b}}$.*** Let $\mathfrak{u}(\mathfrak{b}) \subset \mathfrak{u}$ and $\mathfrak{u}_{\mathfrak{m}} \subset \mathfrak{u}$ be the compact forms of $\mathfrak{z}(\mathfrak{b})$ and $\mathfrak{m}$. Then,

$$\mathfrak{u}(\mathfrak{b}) = \sqrt{-1}\mathfrak{b} \oplus \mathfrak{u}_{\mathfrak{m}}, \quad \mathfrak{u}_{\mathfrak{m}} = \sqrt{-1}\mathfrak{p}_{\mathfrak{m}} \oplus \mathfrak{k}_{\mathfrak{m}}. \tag{6-24}$$

Since $\delta(M)=0$, we know M has compact center. By [Knapp 1986, Proposition 5.3], let U_M be the compact form of M.

Let $U(\mathfrak{b})\subset U$, $A_0\subset U$ be the connected subgroups of U associated with Lie algebras $\mathfrak{u}(\mathfrak{b})$, $\sqrt{-1}\mathfrak{b}$. By (6-24), A_0 is in the center of $U(\mathfrak{b})$, and

$$U(\mathfrak{b})=A_0U_M. \tag{6-25}$$

By [Knapp 2002, Corollary 4.51], the stabilizer of $\mathfrak{b}$ in U is a closed connected subgroup of U, and so it coincides with $U(\mathfrak{b})$.

Proposition 6.6. *The group A_0 is closed in U, and is diffeomorphic to a circle $\mathbb{S}^1$.*

Proof. The proof of Proposition 6.6, based on the classification theory of real simple Lie algebras, will be given in Section 6H. □

Set

$$Y_{\mathfrak{b}}=U/U(\mathfrak{b}). \tag{6-26}$$

We will see that $Y_{\mathfrak{b}}$ is a compact Hermitian symmetric space.

Recall that the bilinear form $-B$ induces an $\mathrm{Ad}(U)$-invariant metric on $\mathfrak{u}$. Let $\mathfrak{u}^{\perp}(\mathfrak{b})$ be the orthogonal space to $\mathfrak{u}(\mathfrak{b})$ in $\mathfrak{u}$ such that

$$\mathfrak{u}=\mathfrak{u}(\mathfrak{b})\oplus\mathfrak{u}^{\perp}(\mathfrak{b}). \tag{6-27}$$

Also, we have

$$\mathfrak{u}^{\perp}(\mathfrak{b})=\sqrt{-1}\mathfrak{p}^{\perp}(\mathfrak{b})\oplus\mathfrak{k}^{\perp}(\mathfrak{b}). \tag{6-28}$$

By (6-12)–(6-14), we have

$$[\mathfrak{u}(\mathfrak{b}),\mathfrak{u}(\mathfrak{b})]\subset\mathfrak{u}(\mathfrak{b}),\quad [\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^{\perp}(\mathfrak{b})]\subset\mathfrak{u}^{\perp}(\mathfrak{b}),\quad [\mathfrak{u}^{\perp}(\mathfrak{b}),\mathfrak{u}^{\perp}(\mathfrak{b})]\subset\mathfrak{u}(\mathfrak{b}). \tag{6-29}$$

Thus, $(\mathfrak{u},\mathfrak{u}(\mathfrak{b}))$ is a symmetric pair.

Set

$$J=\sqrt{-1}\,\mathrm{ad}(a_0)|_{\mathfrak{u}^{\perp}(\mathfrak{b})}\in\mathrm{End}(\mathfrak{u}^{\perp}(\mathfrak{b})). \tag{6-30}$$

By (6-7)–(6-10), J is a $U(\mathfrak{b})$-invariant complex structure on $\mathfrak{u}^{\perp}(\mathfrak{b})$ which preserves the restriction $B|_{\mathfrak{u}^{\perp}(\mathfrak{b})}$. Moreover, $\mathfrak{n}_{\mathbb{C}}=\mathfrak{n}\otimes_{\mathbb{R}}\mathbb{C}$ and $\bar{\mathfrak{n}}_{\mathbb{C}}=\bar{\mathfrak{n}}\otimes_{\mathbb{R}}\mathbb{C}$ are the eigenspaces of J associated with the eigenvalues $\sqrt{-1}$ and $-\sqrt{-1}$ such that

$$\mathfrak{u}^{\perp}(\mathfrak{b})\otimes_{\mathbb{R}}\mathbb{C}=\mathfrak{n}_{\mathbb{C}}\oplus\bar{\mathfrak{n}}_{\mathbb{C}}. \tag{6-31}$$

The bilinear form $-B$ induces a Hermitian metric on $\mathfrak{n}_{\mathbb{C}}$ such that, for $f_1,f_2\in\mathfrak{n}_{\mathbb{C}}$,

$$\langle f_1,f_2\rangle_{\mathfrak{n}_{\mathbb{C}}}=-B(f_1,\bar{f}_2). \tag{6-32}$$

Since J commutes with the action of $U(\mathfrak{b})$, we know $U(\mathfrak{b})$ preserves the splitting (6-31). Therefore, $U(\mathfrak{b})$ acts on $\mathfrak{n}_{\mathbb{C}}$ and $\bar{\mathfrak{n}}_{\mathbb{C}}$. In particular, $U(\mathfrak{b})$ acts on $\Lambda^{\bullet}(\bar{\mathfrak{n}}_{\mathbb{C}}^{*})$. If $S^{\mathfrak{u}^{\perp}(\mathfrak{b})}$ is the spinor of $(\mathfrak{u}^{\perp}(\mathfrak{b}),-B)$, by [Hitchin 1974], we have the isomorphism of representations of $U(\mathfrak{b})$,

$$\Lambda^{\bullet}(\bar{\mathfrak{n}}_{\mathbb{C}}^{*})\simeq S^{\mathfrak{u}^{\perp}(\mathfrak{b})}\otimes\det(\mathfrak{n}_{\mathbb{C}})^{\frac{1}{2}}. \tag{6-33}$$

Note that M has compact center Z_M. By [Knapp 1986, Proposition 5.5], M is a product of a connected semisimple Lie group and the connected component of the identity in Z_M. Since both of these two groups act trivially on $\det(\mathfrak{n})$, the same is true for M. Since the action of U_M on $\mathfrak{n}_{\mathbb{C}}$ can be obtained by the restriction of the induced action of $M_{\mathbb{C}}$ on $\mathfrak{n}_{\mathbb{C}}$, we know U_M acts trivially on $\det(\mathfrak{n}_{\mathbb{C}})$. By (6-33), we have the isomorphism of representations of U_M,

$$\Lambda^\bullet(\bar{\mathfrak{n}}_{\mathbb{C}}^*) \simeq S^{\mathfrak{u}^\perp(\mathfrak{b})}. \tag{6-34}$$

As in Section 4A, let $\omega^{\mathfrak{u}}$ be the canonical left invariant 1-form on U with values in $\mathfrak{u}$, and let $\omega^{\mathfrak{u}(\mathfrak{b})}$ and $\omega^{\mathfrak{u}^\perp(\mathfrak{b})}$ be the $\mathfrak{u}(\mathfrak{b})$ and $\mathfrak{u}^\perp(\mathfrak{b})$ components of $\omega^{\mathfrak{u}}$, so that

$$\omega^{\mathfrak{u}} = \omega^{\mathfrak{u}(\mathfrak{b})} + \omega^{\mathfrak{u}^\perp(\mathfrak{b})}. \tag{6-35}$$

Then, $U \to Y_{\mathfrak{b}}$ is a $U(\mathfrak{b})$-principle bundle, equipped with a connection form $\omega^{\mathfrak{u}(\mathfrak{b})}$. Let $\Omega^{\mathfrak{u}(\mathfrak{b})}$ be the curvature form. As in (4-3), we have

$$\Omega^{\mathfrak{u}(\mathfrak{b})} = -\tfrac{1}{2}[\omega^{\mathfrak{u}^\perp(\mathfrak{b})}, \omega^{\mathfrak{u}^\perp(\mathfrak{b})}]. \tag{6-36}$$

The real tangent bundle

$$TY_{\mathfrak{b}} = U \times_{U(\mathfrak{b})} \mathfrak{u}^\perp(\mathfrak{b}) \tag{6-37}$$

is equipped with a Euclidean metric and a Euclidean connection $\nabla^{TY_{\mathfrak{b}}}$, which coincides with the Levi-Civita connection. By (6-30), J induces an almost complex structure on $TY_{\mathfrak{b}}$. Let $T^{(1,0)}Y_{\mathfrak{b}}$ and $T^{(0,1)}Y_{\mathfrak{b}}$ be the holomorphic and antiholomorphic tangent bundles. Then

$$T^{(1,0)}Y_{\mathfrak{b}} = U \times_{U(\mathfrak{b})} \mathfrak{n}_{\mathbb{C}}, \quad T^{(0,1)}Y_{\mathfrak{b}} = U \times_{U(\mathfrak{b})} \bar{\mathfrak{n}}_{\mathbb{C}}. \tag{6-38}$$

By (6-9) and (6-38), J is integrable.

The form $-B(\cdot, J\cdot)$ induces a Kähler form $\omega^{Y_{\mathfrak{b}}} \in \Omega^2(Y_{\mathfrak{b}})$ on $Y_{\mathfrak{b}}$. Clearly, $\omega^{Y_{\mathfrak{b}}}$ is closed, and therefore $(Y_{\mathfrak{b}}, \omega^{Y_{\mathfrak{b}}})$ is a Kähler manifold. Let $f_1, \dots, f_{2l} \in \mathfrak{n}$ be such that

$$-B(f_i, \bar{f}_j) = \delta_{ij}. \tag{6-39}$$

Then $f_1, \dots, f_{2l}$ is an orthogonal basis of $\mathfrak{n}_{\mathbb{C}}$ with respect to $\langle \cdot, \cdot \rangle_{\mathfrak{n}_{\mathbb{C}}}$. Let $f^1, \dots, f^{2l}$ be the dual base of $\mathfrak{n}_{\mathbb{C}}^*$. The Kähler form $\omega^{Y_{\mathfrak{b}}}$ on $Y_{\mathfrak{b}}$ is given by

$$\omega^{Y_{\mathfrak{b}}} = -\sum_{1 \leqslant i,j \leqslant 2l} B(f_i, J\bar{f}_j) f^i \bar{f}^j = -\sqrt{-1} \sum_{1 \leqslant i \leqslant 2l} f^i \bar{f}^i. \tag{6-40}$$

Let us give a more explicit description of $Y_{\mathfrak{b}}$, although this description will not be needed in the following sections.

Proposition 6.7. *The homogenous space $Y_{\mathfrak{b}}$ is an irreducible compact Hermitian symmetric space of type AIII or BDI.*

Proof. The proof of Proposition 6.7, based on the classification theory of real simple Lie algebras, will be given in Section 6J. □

Since $U_{\mathfrak{m}}$ acts on $\mathfrak{u}_{\mathfrak{m}}$ and A_0 acts trivially on $\mathfrak{u}_{\mathfrak{m}}$, by (6-25), we have $U(\mathfrak{b})$ acts on $\mathfrak{u}_{\mathfrak{m}}$. Put

$$N_{\mathfrak{b}} = U \times_{U(\mathfrak{b})} \mathfrak{u}_{\mathfrak{m}}. \tag{6-41}$$

Then, $N_{\mathfrak{b}}$ is a Euclidean vector bundle on $Y_{\mathfrak{b}}$ equipped with a metric connection $\nabla^{N_{\mathfrak{b}}}$. We equip the trivial connection $\nabla^{\sqrt{-1}\mathfrak{b}}$ on the trivial line bundle $\sqrt{-1}\mathfrak{b}$ on $Y_{\mathfrak{b}}$. Since $U(\mathfrak{b})$ preserves the first splitting in (6-24), we have

$$\sqrt{-1}\mathfrak{b} \oplus N_{\mathfrak{b}} = U \times_{U_{\mathfrak{b}}} U(\mathfrak{b}). \tag{6-42}$$

Moreover, the induced connection is given by

$$\nabla^{\sqrt{-1}\mathfrak{b}\oplus N_{\mathfrak{b}}} = \nabla^{\sqrt{-1}\mathfrak{b}} \oplus \nabla^{N_{\mathfrak{b}}}. \tag{6-43}$$

By (6-27), (6-37), and (6-42), we have

$$TY_{\mathfrak{b}} \oplus \sqrt{-1}\mathfrak{b} \oplus N_{\mathfrak{b}} = \mathfrak{u}, \tag{6-44}$$

where $\mathfrak{u}$ stands for the corresponding trivial bundle on $Y_{\mathfrak{b}}$.

Proposition 6.8. *The following identity of closed forms holds on* $Y_{\mathfrak{b}}$:

$$\widehat{A}(TY_{\mathfrak{b}}, \nabla^{TY_{\mathfrak{b}}})\widehat{A}(N_{\mathfrak{b}}, \nabla^{N_{\mathfrak{b}}}) = 1. \tag{6-45}$$

Proof. Proceeding as in [Bismut 2011, Proposition 7.1.1], by (6-27), (6-37), and (6-42), we have

$$\widehat{A}(TY_{\mathfrak{b}}, \nabla^{TY_{\mathfrak{b}}})\widehat{A}\big(\sqrt{-1}\mathfrak{b} \oplus N_{\mathfrak{b}}, \nabla^{\sqrt{-1}\mathfrak{b}\oplus N_{\mathfrak{b}}}\big) = 1. \tag{6-46}$$

By (6-43), we have

$$\widehat{A}\big(\sqrt{-1}\mathfrak{b} \oplus N_{\mathfrak{b}}, \nabla^{\sqrt{-1}\mathfrak{b}\oplus N_{\mathfrak{b}}}\big) = \widehat{A}(N_{\mathfrak{b}}, \nabla^{N_{\mathfrak{b}}}). \tag{6-47}$$

By (6-46) and (6-47), we get (6-45). □

Recall that the curvature form $\Omega^{\mathfrak{u}(\mathfrak{b})}$ is a 2-form on $Y_{\mathfrak{b}}$ with values in $U \times_{U(\mathfrak{b})} \mathfrak{u}(\mathfrak{b})$. Recall that $a_0 \in \mathfrak{b}$ is defined in (6-8). Let $\Omega^{\mathfrak{u}_{\mathfrak{m}}}$ be the $\mathfrak{u}_{\mathfrak{m}}$-component of $\Omega^{\mathfrak{u}(\mathfrak{b})}$. By (6-8), (6-36) and (6-40), we have

$$\Omega^{\mathfrak{u}(\mathfrak{b})} = \sqrt{-1}\frac{a_0}{|a_0|^2} \otimes \omega^{Y_{\mathfrak{b}}} + \Omega^{\mathfrak{u}_{\mathfrak{m}}}. \tag{6-48}$$

By (6-48), the curvature of $(N_{\mathfrak{b}}, \nabla^{N_{\mathfrak{b}}})$ is given by

$$R^{N_{\mathfrak{b}}} = \mathrm{ad}(\Omega^{\mathfrak{u}(\mathfrak{b})})|_{\mathfrak{u}_{\mathfrak{m}}} = \mathrm{ad}(\Omega^{\mathfrak{u}_{\mathfrak{m}}})|_{\mathfrak{u}_{\mathfrak{m}}}. \tag{6-49}$$

Also, $B(\Omega^{\mathfrak{u}(\mathfrak{b})}, \Omega^{\mathfrak{u}(\mathfrak{b})})$ and $B(\Omega^{\mathfrak{u}_{\mathfrak{m}}}, \Omega^{\mathfrak{u}_{\mathfrak{m}}})$ are well defined 4-forms on $Y_{\mathfrak{b}}$. We have an analogue of [Bismut 2011, Equation (7.5.19)].

Proposition 6.9. *The following identities hold*:

$$B(\Omega^{\mathfrak{u}(\mathfrak{b})}, \Omega^{\mathfrak{u}(\mathfrak{b})}) = 0, \quad B(\Omega^{\mathfrak{u}_{\mathfrak{m}}}, \Omega^{\mathfrak{u}_{\mathfrak{m}}}) = \frac{\omega^{Y_{\mathfrak{b}},2}}{|a_0|^2}. \tag{6-50}$$

Proof. If $e_1, \dots, e_{4l}$ is an orthogonal basis of $\mathfrak{u}^{\perp}(\mathfrak{b})$, by (6-36), we have

$$\begin{aligned} B(\Omega^{\mathfrak{u}(\mathfrak{b})}, \Omega^{\mathfrak{u}(\mathfrak{b})}) &= \frac{1}{4} \sum_{1 \leqslant i,j,i',j' \leqslant 4l} B\big([e_i, e_j], [e_{i'}, e_{j'}]\big) e^i e^j e^{i'} e^{j'} \\ &= \frac{1}{4} \sum_{1 \leqslant i,j,i',j' \leqslant 4l} B\big([[e_i, e_j], e_{i'}], e_{j'}\big) e^i e^j e^{i'} e^{j'}. \end{aligned} \tag{6-51}$$

Using the Jacobi identity and (6-51), we get the first equation of (6-50).

The second equation of (6-50) is a consequence of (6-48) and the first equation of (6-50). □

6C. *Auxiliary virtual representations of K.* Let $\mathrm{RO}(K_M)$ and $\mathrm{RO}(K)$ be the real representation rings of K_M and K. Let $\iota : K_M \to K$ be the injection. We denote by

$$\iota^* : \mathrm{RO}(K) \to \mathrm{RO}(K_M) \tag{6-52}$$

the induced morphism of rings. Since K_M and K have the same maximal torus T, we know ι^* is injective.

Proposition 6.10. *The following identity in* $\mathrm{RO}(K_M)$ *holds*:

$$\iota^*\bigg(\sum_{i=1}^{m} (-1)^{i-1} i \Lambda^i(\mathfrak{p}^*)\bigg) = \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} \sum_{j=0}^{2l} (-1)^{i+j} \Lambda^i(\mathfrak{p}_{\mathfrak{m}}^*) \otimes \Lambda^j(\mathfrak{n}^*). \tag{6-53}$$

Proof. For a representation V of K_M, we use the multiplication notation introduced by Hirzebruch. Put

$$\Lambda_y(V) = \sum_i y^i \Lambda^i(V), \tag{6-54}$$

a polynomial of y with coefficients in $\mathrm{RO}(K_M)$. In particular,

$$\Lambda_{-1}(V) = \sum_i (-1)^i \Lambda^i(V), \quad \Lambda'_{-1}(V) = \sum_i (-1)^{i-1} i \Lambda^i(V). \tag{6-55}$$

Denote by $\mathbf{1}$ the trivial representation. Since $\Lambda_1(\mathbf{1}) = 0$ and $\Lambda'_{-1}(\mathbf{1}) = \mathbf{1}$, we get

$$\Lambda'_{-1}(V \oplus \mathbf{1}) = \Lambda_{-1}(V). \tag{6-56}$$

By (6-2), (6-5), and the fact that K_M acts trivially on $\mathfrak{b}$, we have the isomorphism of K_M-representations

$$\mathfrak{p} \simeq \mathbf{1} \oplus \mathfrak{p}_{\mathfrak{m}} \oplus \mathfrak{n}. \tag{6-57}$$

Taking $V = \mathfrak{p}_{\mathfrak{m}} \oplus \mathfrak{n}$, by (6-56) and (6-57), we get (6-53). □

The following theorem is crucial.

Theorem 6.11. *The adjoint representation of* K_M *on* $\mathfrak{n}$ *has a unique lift in* $\mathrm{RO}(K)$.

Proof. The injectivity of ι^* implies the uniqueness. The proof of the existence of the lifting of $\mathfrak{n}$, based on the classification theorem of real simple Lie algebras, will be given in Section 6I. □

Corollary 6.12. *For* $i, j \in \mathbb{N}$, *the adjoint representations of* K_M *on* $\Lambda^i(\mathfrak{p}_{\mathfrak{m}}^*)$ *and* $\Lambda^j(\mathfrak{n}^*)$ *have unique lifts in* $\mathrm{RO}(K)$.

Proof. As before, it is enough to show the existence of lifts. Since the representation of K_M on $\mathfrak{n}$ lifts to K, the same is true for the $\Lambda^j(\mathfrak{n}^*)$. By (6-57), this extends to the $\Lambda^i(\mathfrak{p}_\mathfrak{m}^*)$. □

Denote by η_j the adjoint representation of M on $\Lambda^j(\mathfrak{n}^*)$. Recall that by (6-31), $U(\mathfrak{b})$ acts on $\mathfrak{n}_\mathbb{C}$. Recall also that $C^{\mathfrak{u}_\mathfrak{m},\eta_j} \in \mathrm{End}(\Lambda^j(\mathfrak{n}_\mathbb{C}^*))$, $C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})} \in \mathrm{End}(\mathfrak{u}^\perp(\mathfrak{b}))$ are defined in (3-8).

Proposition 6.13. *For $0 \leqslant j \leqslant 2l$, the operator $C^{\mathfrak{u}_\mathfrak{m},\eta_j}$ is a scalar such that*

$$C^{\mathfrak{u}_\mathfrak{m},\eta_j} = \tfrac{1}{8}\mathrm{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}] + (j-l)^2|\alpha|^2. \tag{6-58}$$

Proof. Equation (6-58) was proved in [Moscovici and Stanton 1991, Lemma 2.5]. We give here a more conceptual proof.

Recall that $(\mathfrak{u},\mathfrak{u}(\mathfrak{b}))$ is a compact symmetric pair. Let $S^{\mathfrak{u}^\perp(\mathfrak{b})}$ be the $\mathfrak{u}^\perp(\mathfrak{b})$-spinors [Bismut 2011, Section 7.2]. Let $C^{\mathfrak{u}(\mathfrak{b}),S^{\mathfrak{u}^\perp(\mathfrak{b})}}$ be the Casimir element of $\mathfrak{u}(\mathfrak{b})$ acting on $S^{\mathfrak{u}^\perp(\mathfrak{b})}$ defined as in (3-8). By (7.8.6) of the same paper, $C^{\mathfrak{u}(\mathfrak{b}),S^{\mathfrak{u}^\perp(\mathfrak{b})}}$ is a scalar such that

$$C^{\mathfrak{u}(\mathfrak{b}),S^{\mathfrak{u}^\perp(\mathfrak{b})}} = \tfrac{1}{8}\mathrm{Tr}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}]. \tag{6-59}$$

Let $C^{\mathfrak{u}_\mathfrak{m},\Lambda^\bullet(\bar{\mathfrak{n}}_\mathbb{C}^*)}$ be the Casimir element of $\mathfrak{u}_\mathfrak{m}$ acting on $\Lambda^\bullet(\bar{\mathfrak{n}}_\mathbb{C}^*)$. By (3-7), (6-33) and (6-34), we have

$$C^{\mathfrak{u}(\mathfrak{b}),S^{\mathfrak{u}^\perp(\mathfrak{b})}} = C^{\mathfrak{u}_\mathfrak{m},\Lambda^\bullet(\bar{\mathfrak{n}}_\mathbb{C}^*)} - (\mathrm{Ad}(a_1)|_{\Lambda^\bullet(\bar{\mathfrak{n}}_\mathbb{C}^*)\otimes\det^{-1/2}(\mathfrak{n}_\mathbb{C})})^2. \tag{6-60}$$

By (6-7), we have

$$\mathrm{Ad}(a_1)|_{\Lambda^j(\bar{\mathfrak{n}}_\mathbb{C}^*)\otimes\det^{-1/2}(\mathfrak{n}_\mathbb{C})} = (j-l)|\alpha|. \tag{6-61}$$

By (6-59)–(6-61), we get (6-58). □

Let $\gamma = e^a k^{-1} \in G$ be such that (3-9) holds. Since $\Lambda^\bullet(\mathfrak{p}_\mathfrak{m}^*) \in \mathrm{RO}(K)$, for $Y \in \mathfrak{k}(\gamma)$, we know ${\mathrm{Tr_s}}^{\Lambda^\bullet(\mathfrak{p}_\mathfrak{m}^*)}[k^{-1}\exp(-iY)]$ is well defined. We have an analogue of (4-44).

Proposition 6.14. *If* $\dim\mathfrak{b}(\gamma) \geqslant 2$, *then for* $Y \in \mathfrak{k}(\gamma)$, *we have*

$${\mathrm{Tr_s}}^{\Lambda^\bullet(\mathfrak{p}_\mathfrak{m}^*)}[k^{-1}\exp(-iY)] = 0. \tag{6-62}$$

Proof. The proof of Proposition 6.14, based on the classification theory of real simple Lie algebras, will be given in Section 6L. □

6D. ***A classification of real reductive Lie algebra $\mathfrak{g}$ with $\delta(\mathfrak{g}) = 1$.*** Recall that G is a real reductive group with compact center such that $\delta(G) = 1$.

Theorem 6.15. *We have a decomposition of Lie algebras*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2, \tag{6-63}$$

where

$$\mathfrak{g}_1 = \mathfrak{sl}_3(\mathbb{R}) \quad or \quad \mathfrak{so}(p,q), \tag{6-64}$$

with $pq > 1$ odd, and $\mathfrak{g}_2$ is real reductive with $\delta(\mathfrak{g}_2) = 0$.

Proof. Since G has compact center, by (3-6), $\mathfrak{z}_{\mathfrak{p}} = 0$. By (3-25), we have

$$\delta([\mathfrak{g}, \mathfrak{g}]) = 1. \tag{6-65}$$

As in [Bismut 2011, Remark 7.9.2], by the classification theory of real simple Lie algebras, we have

$$[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}_1 \oplus \mathfrak{g}_2', \tag{6-66}$$

where

$$\mathfrak{g}_1 = \mathfrak{sl}_3(\mathbb{R}) \quad \text{or} \quad \mathfrak{so}(p, q), \tag{6-67}$$

with $pq > 1$ odd, and where $\mathfrak{g}_2'$ is semisimple with $\delta(\mathfrak{g}_2') = 0$. Take

$$\mathfrak{g}_2 = \mathfrak{z}_{\mathfrak{k}} \oplus \mathfrak{g}_2'. \tag{6-68}$$

By (3-24), (6-66)–(6-68), we get (6-63). □

6E. *The group* $\mathbf{SL_3(\mathbb{R})}$. In this subsection, we assume that $G = \mathrm{SL}_3(\mathbb{R})$, so that $K = \mathrm{SO}(3)$. We have

$$\mathfrak{p} = \left\{ \begin{pmatrix} x & a_1 & a_2 \\ a_1 & y & a_3 \\ a_2 & a_3 & -x-y \end{pmatrix} : x, y, a_1, a_2, a_3 \in \mathbb{R} \right\}, \quad \mathfrak{k} = \left\{ \begin{pmatrix} 0 & a_1 & a_2 \\ -a_1 & 0 & a_3 \\ -a_2 & -a_3 & 0 \end{pmatrix} : a_1, a_2, a_3 \in \mathbb{R} \right\}. \tag{6-69}$$

Let

$$T = \left\{ \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} : A \in \mathrm{SO}(2) \right\} \subset K \tag{6-70}$$

be a maximal torus of K.

By (3-18), (6-69) and (6-70), we have

$$\mathfrak{b} = \left\{ \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & -2x \end{pmatrix} : x \in \mathbb{R} \right\} \subset \mathfrak{p}. \tag{6-71}$$

By (6-71), we get

$$\mathfrak{p}_{\mathfrak{m}} = \left\{ \begin{pmatrix} x & a_1 & 0 \\ a_1 & -x & 0 \\ 0 & 0 & 0 \end{pmatrix} : x, a_1 \in \mathbb{R} \right\}, \quad \mathfrak{p}^{\perp}(\mathfrak{b}) = \left\{ \begin{pmatrix} 0 & 0 & a_2 \\ 0 & 0 & a_3 \\ a_2 & a_3 & 0 \end{pmatrix} : a_2, a_3 \in \mathbb{R} \right\}. \tag{6-72}$$

Also,

$$\mathfrak{k}_{\mathfrak{m}} = \mathfrak{t}, \quad K_M = T, \quad M = \left\{ \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} : A \in \mathrm{SL}_2(\mathbb{R}) \right\}. \tag{6-73}$$

By (6-71), we can orient $\mathfrak{b}$ by $x > 0$. Thus,

$$\mathfrak{n} = \left\{ \begin{pmatrix} 0 & 0 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix} : a_2, a_3 \in \mathbb{R} \right\}. \tag{6-74}$$

By (6-71) and (6-74), since for $x \in \mathbb{R},\ a_2 \in \mathbb{R},\ a_3 \in \mathbb{R}$,

$$\left[\begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & -2x \end{pmatrix}, \begin{pmatrix} 0 & 0 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix}\right] = 3x \begin{pmatrix} 0 & 0 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6-75}$$

we find that $\mathfrak{b}$ acts on $\mathfrak{n}$ as a scalar.

Denote by $\mathrm{Isom}^0(\mathrm{G/K})$ the connected component of the identity of the isometric group of $X = G/K$. Since G acts isometrically on G/K, we have the morphism of groups

$$i_G : G \to \mathrm{Isom}^0(\mathrm{G/K}). \tag{6-76}$$

Proposition 6.16. *The morphism i_G is an isomorphism*; *i.e.*,

$$\mathrm{SL}_3(\mathbb{R}) \simeq \mathrm{Isom}^0(\mathrm{SL}_3(\mathbb{R})/\,\mathrm{SO}(3)). \tag{6-77}$$

Proof. By [Helgason 1978, Theorem V.4.1], it is enough to show that K acts on $\mathfrak{p}$ effectively. Assume that $k \in K$ acts on $\mathfrak{p}$ as the identity. Thus, k fixes the elements of $\mathfrak{b}$. As in (6-73), there is $A \in \mathrm{GL}_2(\mathbb{R})$ such that

$$k = \begin{pmatrix} A & 0 \\ 0 & \det^{-1}(A) \end{pmatrix}. \tag{6-78}$$

Since k fixes also the elements of $\mathfrak{p}^{\perp}(\mathfrak{b})$, by (6-72) and (6-78), we get $A = 1$. Therefore, $k = 1$. □

6F. *The group $G = \mathrm{SO}^0(p,q)$ with $pq > 1$ odd.* In this subsection, we assume that $G = \mathrm{SO}^0(p,q)$, so that $K = \mathrm{SO}(p) \times \mathrm{SO}(q)$, with $pq > 1$ odd.

In the sequel, if $l, l' \in \mathbb{N}^*$, let $\mathrm{Mat}_{l,l'}(\mathbb{R})$ be the space of real matrices of l rows and l' columns. If $L \subset \mathrm{Mat}_{l,l}(\mathbb{R})$ is a matrix group, we denote by σ_l the standard representation of L on $\mathbb{R}^l$. We have

$$\mathfrak{p} = \left\{\begin{pmatrix} 0 & B \\ B^t & 0 \end{pmatrix} : B \in \mathrm{Mat}_{p,q}(\mathbb{R})\right\}, \quad \mathfrak{k} = \left\{\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} : A \in \mathfrak{so}(p),\ D \in \mathfrak{so}(q)\right\}. \tag{6-79}$$

Let

$$T_{p-1} = \left\{\begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_{(p-1)/2} \end{pmatrix} : A_1, \dots, A_{(p-1)/2} \in \mathrm{SO}(2)\right\} \subset \mathrm{SO}(p-1) \tag{6-80}$$

be a maximal torus of $\mathrm{SO}(p-1)$. Then,

$$T = \left\{\begin{pmatrix} A & 0 & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & B \end{pmatrix} \in K : A \in T_{p-1},\ B \in T_{q-1}\right\} \subset K \tag{6-81}$$

is a maximal torus of K.

By (3-18) and (6-81), we have

$$\begin{aligned}
\mathfrak{b} &= \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{p} : x \in \mathbb{R} \right\}, \\
\mathfrak{p}_{\mathfrak{m}} &= \left\{ \begin{pmatrix} 0 & 0 & B \\ 0 & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & 0 \\ B^t & 0 & 0 \end{pmatrix} \in \mathfrak{p} : B \in \mathrm{Mat}_{p-1,q-1}(\mathbb{R}) \right\}, \\
\mathfrak{p}^{\perp}(\mathfrak{b}) &= \left\{ \begin{pmatrix} 0 & 0 \ v_1 & 0 \\ 0 & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & v_2^t \\ v_1^t & & 0 \\ 0 & v_2 \ 0 & 0 \end{pmatrix} \in \mathfrak{p} : v_1 \in \mathbb{R}^{p-1},\ v_2 \in \mathbb{R}^{q-1} \right\},
\end{aligned} \tag{6-82}$$

where v_1, v_2 are considered as column vectors. Also,

$$\mathfrak{k}_{\mathfrak{m}} = \left\{ \begin{pmatrix} A & 0 & 0 \\ 0 & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & 0 \\ 0 & 0 & D \end{pmatrix} \in \mathfrak{k} : A \in \mathfrak{so}(p-1),\ D \in \mathfrak{so}(q-1) \right\}. \tag{6-83}$$

By (6-82) and (6-83), we get

$$\begin{aligned}
M &= \left\{ \begin{pmatrix} A & 0 & B \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0 \\ C & 0 & D \end{pmatrix} \in G : \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{SO}^0(p-1, q-1) \right\}, \\
K_M &= \left\{ \begin{pmatrix} A & 0 & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & D \end{pmatrix} \in K : A \in \mathrm{SO}(p-1), D \in \mathrm{SO}(q-1) \right\}.
\end{aligned} \tag{6-84}$$

By (6-82), we can orient $\mathfrak{b}$ by $x > 0$. Then,

$$\mathfrak{n} = \left\{ \begin{pmatrix} 0 & -v_1 \ v_1 & 0 \\ v_1^t & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & v_2^t \\ v_1^t & & v_2^t \\ 0 & v_2 \ -v_2 & 0 \end{pmatrix} \in \mathfrak{g} : v_1 \in \mathbb{R}^{p-1},\ v_2 \in \mathbb{R}^{q-1} \right\}. \tag{6-85}$$

By (6-82) and (6-85), since for $x \in \mathbb{R},\ v_1 \in \mathbb{R}^{p-1},\ v_2 \in \mathbb{R}^{q-1}$,

$$\left[\begin{pmatrix} 0 & 0 & 0 \\ 0 & \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -v_1 \ v_1 & 0 \\ v_1^t & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & v_2^t \\ v_1^t & & v_2^t \\ 0 & v_2 \ -v_2 & 0 \end{pmatrix} \right] = x \begin{pmatrix} 0 & -v_1 \ v_1 & 0 \\ v_1^t & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & v_2^t \\ v_1^t & & v_2^t \\ 0 & v_2 \ -v_2 & 0 \end{pmatrix}, \tag{6-86}$$

we find that $\mathfrak{b}$ acts on $\mathfrak{n}$ as a scalar.

Proposition 6.17. *We have the isomorphism of Lie groups*

$$\mathrm{SO}^0(p,q) \simeq \mathrm{Isom}^0\big(\mathrm{SO}^0(p,q)/\,\mathrm{SO}(p)\times \mathrm{SO}(q)\big), \tag{6-87}$$

where $pq > 1$ is odd.

Proof. As in the proof of Proposition 6.16, it is enough to show that K acts effectively on $\mathfrak{p}$. The representation of $K \simeq \mathrm{SO}(p)\times\mathrm{SO}(q)$ on $\mathfrak{p}$ is equivalent to $\sigma_p \boxtimes \sigma_q$. Assume that $(k_1,k_2)\in \mathrm{SO}(p)\times\mathrm{SO}(q)$ acts on $\mathbb{R}^p \boxtimes \mathbb{R}^q$ as the identity. If λ is any eigenvalue of k_1 and if μ is any eigenvalue of k_2, then

$$\lambda\mu = 1. \tag{6-88}$$

By (6-88), both k_1 and k_2 are scalars. Using the fact that $\det(k_1) = \det(k_2) = 1$ and that p, q are odd, we deduce $k_1 = 1$ and $k_2 = 1$. □

6G. ***The isometry group of X.*** We return to the general case, where G is only assumed to have compact center and be such that $\delta(G) = 1$.

Proposition 6.18. *The symmetric space G/K is of noncompact type.*

Proof. Let Z_G^0 be the connected component of the identity in Z_G, and let $G_{ss} \subset G$ be the connected subgroup of G associated with the Lie algebra $[\mathfrak{g},\mathfrak{g}]$. By [Knapp 1986, Proposition 5.5], G_{ss} is closed in G such that

$$G = Z_G^0 G_{ss}. \tag{6-89}$$

Moreover, G_{ss} is semisimple with finite center, with maximal compact subgroup $K_{ss} = G_{ss} \cap K$. Also, the imbedding $G_{ss} \to G$ induces the diffeomorphism

$$G_{ss}/K_{ss} \simeq G/K. \tag{6-90}$$

Therefore, X is a symmetric space of noncompact type. □

Put

$$G_* = \mathrm{Isom}^0(X), \tag{6-91}$$

and let $K_* \subset G_*$ be the stabilizer of $p1 \in X$ fixed. Then G_* is a semisimple Lie group with trivial center, and with maximal compact subgroup K_*. We denote by $\mathfrak{g}_*$ and $\mathfrak{k}_*$ the Lie algebras of G_* and K_*. Let

$$\mathfrak{g}_* = \mathfrak{p}_* \oplus \mathfrak{k}_* \tag{6-92}$$

be the corresponding Cartan decomposition. Clearly,

$$G_*/K_* \simeq X. \tag{6-93}$$

The morphism $i_G : G \to G_*$ defined in (6-76) induces a morphism $i_{\mathfrak{g}} : \mathfrak{g} \to \mathfrak{g}_*$ of Lie algebras. By (3-4) and (6-93), $i_{\mathfrak{g}}$ induces an isomorphism of vector spaces

$$\mathfrak{p} \simeq \mathfrak{p}_*. \tag{6-94}$$

By the property of $\mathfrak{k}_*$ and by (6-94), we have

$$\mathfrak{k}_* = [\mathfrak{p}_*,\mathfrak{p}_*] = i_{\mathfrak{g}}[\mathfrak{p},\mathfrak{p}] \subset i_{\mathfrak{g}}\mathfrak{k}. \tag{6-95}$$

Thus $i_G, i_{\mathfrak{g}}$ are surjective.

Proposition 6.19. *We have*

$$G_* = G_1 \times G_2, \tag{6-96}$$

where $G_1 = \mathrm{SL}_3(\mathbb{R})$ or $G_1 = \mathrm{SO}^0(p,q)$ with $pq > 1$ odd, and where G_2 is a semisimple Lie group with trivial center with $\delta(G_2) = 0$.

Proof. By [Kobayashi and Nomizu 1963, Theorem IV.6.2], let $X = \prod_{i=1}^{l_1} X_l$ be the de Rham decomposition of (X, g^{TX}). Then every X_i is an irreducible symmetric space of noncompact type. By Theorem VI.3.5 of the same paper, we have

$$G_* = \prod_{i=1}^{l_1} \mathrm{Isom}^0(X_i). \tag{6-97}$$

By Theorem 6.15, (6-77), (6-87) and (6-97), Proposition 6.19 follows. □

6H. *Proof of Proposition 6.2.* By (6-63) and by the definitions of $\mathfrak{b}$ and $\mathfrak{n}$, we have

$$\mathfrak{b}, \mathfrak{n} \subset \mathfrak{g}_1. \tag{6-98}$$

Proposition 6.2 follows from (6-75) and (6-86). □

6I. *Proof of Theorem 6.11.* *The case $G = \mathrm{SL}_3(\mathbb{R})$.* By (6-73) and (6-74), the representation of $K_M \simeq \mathrm{SO}(2)$ on $\mathfrak{n}$ is just σ_2. Note that $K = \mathrm{SO}(3)$. We have the identity in $\mathrm{RO}(K_M)$

$$\iota^*(\sigma_3 - \mathbf{1}) = \sigma_2, \tag{6-99}$$

which says $\mathfrak{n}$ lifts to K.

The case $G = \mathrm{SO}^0(p,q)$ with $pq > 1$ odd. By (6-84) and (6-85), the representation of $K_M \simeq \mathrm{SO}(p-1) \times \mathrm{SO}(q-1)$ on $\mathfrak{n}$ is just $\sigma_{p-1} \boxtimes \mathbf{1} \oplus \mathbf{1} \boxtimes \sigma_{q-1}$. Note that $K = \mathrm{SO}(p) \times \mathrm{SO}(q)$. We have the identity in $\mathrm{RO}(K_M)$

$$\iota^*\big((\sigma_p - \mathbf{1}) \boxtimes \mathbf{1} \oplus \mathbf{1} \boxtimes (\sigma_q - \mathbf{1})\big) = \sigma_{p-1} \boxtimes \mathbf{1} \oplus \mathbf{1} \boxtimes \sigma_{q-1}, \tag{6-100}$$

which says $\mathfrak{n}$ lifts to K.

The case for G_*. This is a consequence of Proposition 6.19 and (6-98)–(6-100).

The general case. Recall that $i_G : G \to G_*$ is a surjective morphism of Lie groups. Therefore, the restriction $i_K : K \to K_*$ of i_G to K is surjective. By (6-94), we have the identity in $\mathrm{RO}(K)$

$$\mathfrak{p} = i_K^*(\mathfrak{p}_*). \tag{6-101}$$

Set

$$\mathfrak{t}_* = i_{\mathfrak{g}}(\mathfrak{t}) \subset \mathfrak{k}_*. \tag{6-102}$$

Since i_K is surjective, by [Bröcker and tom Dieck 1985, Theorem IV.2.9], $\mathfrak{t}_*$ is a Cartan subalgebra of $\mathfrak{k}_*$.

Let $\mathfrak{b}_* \subset \mathfrak{p}_*$ be the analogue of $\mathfrak{b}$ defined by $\mathfrak{t}_*$. Thus,

$$\dim \mathfrak{b}_* = 1, \quad \mathfrak{b}_* = i_{\mathfrak{g}}(\mathfrak{b}). \tag{6-103}$$

We denote by $K_{*,M}$, $\mathfrak{p}_*^{\perp}(\mathfrak{b}_*)$, $\mathfrak{n}_*$ the analogues of K_M, $\mathfrak{p}^{\perp}(\mathfrak{b})$, $\mathfrak{n}$. By (6-94), $i_{\mathfrak{g}}$ induces an isomorphism of vector spaces

$$\mathfrak{p}^{\perp}(\mathfrak{b}) \simeq \mathfrak{p}_*^{\perp}(\mathfrak{b}_*). \tag{6-104}$$

Let $i_{K_M} : K_M \to K_{*,M}$ be the restriction of i_G to K_M. We have the identity in $\mathrm{RO}(K_M)$

$$\mathfrak{p}^{\perp}(\mathfrak{b}) = i_{K_M}^*(\mathfrak{p}_*^{\perp}(\mathfrak{b}_*)). \tag{6-105}$$

Let $\iota' : K_{*,M} \to K_*$ be the imbedding. Then the diagram

$$\begin{tikzcd} K_M \arrow[r, "\iota"] \arrow[d, "i_{K_M}"] & K \arrow[d, "i_K"] \\ K_{*,M} \arrow[r, "\iota'"] & K_* \end{tikzcd} \tag{6-106}$$

commutes. It was proved in the previous step that there is $E \in \mathrm{RO}(K_*)$ such that the following identity in $\mathrm{RO}(K_{*,M})$ holds:

$$\iota'^*(E) = \mathfrak{n}_*. \tag{6-107}$$

By (6-5) and (6-105)–(6-107), we have the identity in $\mathrm{RO}(K_M)$,

$$\mathfrak{n} = \mathfrak{p}^{\perp}(\mathfrak{b}) = i_{K_M}^*(\mathfrak{p}_*^{\perp}(\mathfrak{b}_*)) = i_{K_M}^*(\mathfrak{n}_*) = i_{K_M}^* \iota'^*(E) = \iota^* i_K^*(E), \tag{6-108}$$

which completes the proof of our theorem. □

6J. *Proof of Proposition 6.7.* If $n \in \mathbb{N}$, consider the following closed subgroups:

$$\begin{aligned} &A \in \mathrm{U}(2) \to \begin{pmatrix} A & 0 \\ 0 & \det^{-1}(A) \end{pmatrix} \in \mathrm{SU}(3), \\ &(A, B) \in \mathrm{SO}(n) \times \mathrm{SO}(2) \to \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathrm{SO}(n+2). \end{aligned} \tag{6-109}$$

We state Proposition 6.7 in a more exact way.

Proposition 6.20. *We have the isomorphism of symmetric spaces*

$$Y_{\mathfrak{b}} \simeq \mathrm{SU}(3)/\mathrm{U}(2) \quad \textit{or} \quad \mathrm{SO}(p+q)/\,\mathrm{SO}(p+q-2) \times \mathrm{SO}(2), \tag{6-110}$$

with $pq > 1$ odd.

Proof. Let U_* and $U_*(\mathfrak{b}_*)$ be the analogues of U and $U(\mathfrak{b})$ when G and $\mathfrak{b}$ are replaced by G_* and $\mathfrak{b}_*$. It is enough to show that

$$Y_{\mathfrak{b}} \simeq U_*/U_*(\mathfrak{b}_*). \tag{6-111}$$

Indeed, by the explicit constructions given in Sections 6E and 6F, by Proposition 6.19, and by (6-109), (6-111), we get (6-110).

Let $Z_U \subset U$ be the center of U, and let Z_U^0 be the connected component of the identity in Z_U. Let $U_{ss} \subset U$ be the connected subgroup of U associated to the Lie algebra $[\mathfrak{u}, \mathfrak{u}] \subset \mathfrak{u}$. By [Knapp 1986, Proposition 4.32], U_{ss} is compact, and $U = U_{ss} Z_U^0$.

Let $U_{ss}(\mathfrak{b})$ be the analogue of $U(\mathfrak{b})$ when U is replaced by U_{ss}. Then $U(\mathfrak{b}) = U_{ss}(\mathfrak{b}) Z_U^0$, and the imbedding $U_{ss} \to U$ induces an isomorphism of homogeneous spaces

$$U_{ss}/U_{ss}(\mathfrak{b}) \simeq U/U(\mathfrak{b}). \tag{6-112}$$

Let $\widetilde{U}_{ss}$ be the universal cover of U_{ss}. Since U_{ss} is semisimple, $\widetilde{U}_{ss}$ is compact. We define $\widetilde{U}_{ss}(\mathfrak{b})$ similarly. The canonical projection $\widetilde{U}_{ss} \to U_{ss}$ induces an isomorphism of homogeneous spaces

$$\widetilde{U}_{ss}/\widetilde{U}_{ss}(\mathfrak{b}) \simeq U_{ss}/U_{ss}(\mathfrak{b}). \tag{6-113}$$

Similarly, since U_* is semisimple, if $\widetilde{U}_*$ is a universal cover of U_*, and if we define $\widetilde{U}_*(\mathfrak{b})$ in the same way, we have

$$\widetilde{U}_*/\widetilde{U}_*(\mathfrak{b}) \simeq U_*/U_*(\mathfrak{b}). \tag{6-114}$$

The surjective morphism of Lie algebras $i_{\mathfrak{g}} : \mathfrak{g} \to \mathfrak{g}_*$ induces a surjective morphism of the compact forms $i_{\mathfrak{u}} : \mathfrak{u} \to \mathfrak{u}_*$. Since $\mathfrak{u}_*$ is semisimple, the restriction of $i_{\mathfrak{u}}$ to $[\mathfrak{u}, \mathfrak{u}]$ is still surjective. It lifts to a surjective morphism of simply connected Lie groups

$$\widetilde{U}_{ss} \to \widetilde{U}_*. \tag{6-115}$$

Since any connected, simply connected, semisimple compact Lie group can be written as a product of connected, simply connected, simple compact Lie groups, we can assume that there is a connected and simply connected semisimple compact Lie group U' such that $\widetilde{U}_{ss} = \widetilde{U}_* \times U'$, and that the morphism (6-115) is the canonical projection. Therefore,

$$\widetilde{U}_{ss}/\widetilde{U}_{ss}(\mathfrak{b}) \simeq \widetilde{U}_*/\widetilde{U}_*(\mathfrak{b}_*). \tag{6-116}$$

From (6-26), (6-112)–(6-114) and (6-116), we get (6-111). □

Remark 6.21. The Hermitian symmetric spaces on the right-hand side of (6-110) are irreducible and respectively of type AIII and type BDI in the classification of Cartan [Helgason 1978, p. 518, Table V].

6K. ***Proof of Proposition 6.6.*** We use the notation in Section 6J. By definition, $A_0 \subset U_{ss}$. Let $\widetilde{A}_0 \subset \widetilde{U}_{ss}$ and $A_{*0} \subset U_*$ be the analogues of A_0 when U is replaced by $\widetilde{U}_{ss}$ and U_*. As in the proof of Proposition 6.7, we can show that $\widetilde{A}_0$ is a finite cover of A_0 and A_{*0}.

On the other hand, by the explicit constructions given in Sections 6E, 6F, and by Proposition 6.19, A_{*0} is a circle $\mathbb{S}^1$. Therefore, both $\widetilde{A}_0, A_0$ are circles.

6L. ***Proof of Proposition 6.14.*** We use the notation in Section 4E. Let $\gamma \in G$ be such that $\dim \mathfrak{b}(\gamma) \geqslant 2$. As in (4-39), we assume that $\gamma = e^a k^{-1}$ is such that

$$\mathfrak{t}(\gamma) \subset \mathfrak{t}, \quad k \in T. \tag{6-117}$$

It is enough to show (6-62) for $Y \in \mathfrak{t}(\gamma)$.

For $Y \in \mathfrak{t}(\gamma)$, since $k^{-1}\exp(-iY) \in T$ and $T \subset K^M$, we have

$$\mathrm{Tr_s}^{\Lambda^\bullet(\mathfrak{p}_\mathfrak{m})}[k^{-1}\exp(-iY)] = \det\big(1-\mathrm{Ad}(k)\exp(i\,\mathrm{ad}(Y))\big)\big|_{\mathfrak{p}_\mathfrak{m}}. \tag{6-118}$$

It is enough to show

$$\dim \mathfrak{b}(\gamma)\cap\mathfrak{p}_\mathfrak{m} \geqslant 1. \tag{6-119}$$

Note that $a \neq 0$, otherwise $\dim \mathfrak{b}(\gamma) = 1$. Let

$$a = a^1 + a^2 + a^3 \in \mathfrak{b}\oplus\mathfrak{p}_\mathfrak{m}\oplus\mathfrak{p}^\perp(\mathfrak{b}). \tag{6-120}$$

Since the decomposition $\mathfrak{b}\oplus\mathfrak{p}_\mathfrak{m}\oplus\mathfrak{p}^\perp(\mathfrak{b})$ is preserved by $\mathrm{ad}(\mathfrak{t})$ and $\mathrm{Ad}(T)$, it is also preserved by $\mathrm{ad}(\mathfrak{t}(\gamma))$ and $\mathrm{Ad}(k)$. Since $a \in \mathfrak{b}(\gamma)$, the a_i, $1 \leqslant i \leqslant 3$, all lie in $\mathfrak{b}(\gamma)$. If $a^2 \neq 0$, we get (6-119). If $a^2 = 0$ and $a^3 = 0$, we have $a \in \mathfrak{b}$. Since $a \neq 0$, we have $\mathfrak{b}(\gamma) = \mathfrak{b}$, which is impossible since $\dim b(\gamma) \geqslant 2$.

It remains to consider the case

$$a^2 = 0, \quad a^3 \neq 0. \tag{6-121}$$

We will follow the steps in the proof of Theorem 6.11.

The case $G = \mathrm{SL}_3(\mathbb{R})$. By (6-70) and (6-72), the representation of $T \simeq \mathrm{SO}(2)$ on $\mathfrak{p}^\perp(\mathfrak{b})$ is equivalent to σ_2. A nontrivial element of T never fixes a^3. Therefore,

$$k = 1. \tag{6-122}$$

Since $a \notin \mathfrak{b}$, we know a does not commute with all the elements of $\mathfrak{t}$. From (6-117), we get

$$\dim \mathfrak{t}(\gamma) < \dim \mathfrak{t} = 1. \tag{6-123}$$

Therefore,

$$\mathfrak{t}(\gamma) = 0. \tag{6-124}$$

By (4-36), (6-122) and (6-124), we see that $\mathfrak{b}(\gamma) = \mathfrak{p}$. Therefore,

$$\dim \mathfrak{b}(\gamma)\cap\mathfrak{p}_\mathfrak{m} = \dim\mathfrak{p}_\mathfrak{m}. \tag{6-125}$$

By (6-72) and (6-125), we get (6-119).

The case $G = \mathrm{SO}^0(p,q)$ with $pq > 1$ odd. By (6-82) and (6-84), the representations of $K_M \simeq \mathrm{SO}(p-1)\times\mathrm{SO}(q-1)$ on $\mathfrak{p}_\mathfrak{m}$ and $\mathfrak{p}^\perp(\mathfrak{b})$ are equivalent to $\sigma_{p-1}\boxtimes\sigma_{q-1}$ and $\sigma_{p-1}\boxtimes\mathbf{1}\oplus\mathbf{1}\boxtimes\sigma_{q-1}$. We identify $a^3 \in \mathfrak{p}^\perp(\mathfrak{b})$ with

$$v^1 + v^2 \in \mathbb{R}^{p-1}\oplus\mathbb{R}^{q-1}. \tag{6-126}$$

Then v^1 and v^2 are fixed by $\mathrm{Ad}(k)$ and commute with $\mathfrak{t}(\gamma)$.

If $v^1 \neq 0$ and $v^2 \neq 0$, by (4-36), the nonzero element $v^1\boxtimes v^2 \in \mathbb{R}^{p-1}\boxtimes\mathbb{R}^{q-1} \simeq \mathfrak{p}_\mathfrak{m}$ is in $\mathfrak{b}(\gamma)$. It implies (6-119).

If $v^2 = 0$, we will show that γ can be conjugated into H by an element of K, which implies $\dim\mathfrak{b}(\gamma) = 1$ and contradicts $\dim\mathfrak{b}(\gamma) \geqslant 2$. (The proof for the case $v^1 = 0$ is similar.) Without loss of generality, assume that there exist $s \in \mathbb{N}$ with $1 \leqslant s \leqslant (p-1)/2$ and nonzero complex numbers $\lambda_s, \dots, \lambda_{(p-1)/2} \in \mathbb{C}$ such that

$$v^1 = (0, \dots, 0, \lambda_s, \dots, \lambda_{(p-1)/2}) \in \mathbb{C}^{(p-1)/2} \simeq \mathbb{R}^{p-1}. \tag{6-127}$$

Then there exists $x \in \mathbb{R}$ such that

$$a = \begin{pmatrix} 0 & 0 \; v^1 & 0 \\ 0 & \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ v^{1t} & & \\ 0 & 0 \;\; 0 & 0 \end{pmatrix} \in \mathfrak{p}. \tag{6-128}$$

By (6-81) and (6-117), there exist $A \in T_{p-1}$ and $D \in T_{q-1}$ such that

$$k = \begin{pmatrix} A & 0 & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & D \end{pmatrix} \in T. \tag{6-129}$$

If we identify $T_{p-1} \simeq U(1)^{(p-1)/2}$, there are $\theta_1, \dots, \theta_{(p-1)/2} \in \mathbb{R}$ such that

$$A = (e^{2\sqrt{-1}\pi\theta_1}, \dots, e^{2\sqrt{-1}\pi\theta_{(p-1)/2}}). \tag{6-130}$$

Since k fixes a, by (6-127)–(6-130), for $i = s, \dots, (p-1)/2$, we have

$$e^{2\sqrt{-1}\pi\theta_i} = 1. \tag{6-131}$$

If $W \in \mathfrak{so}(p-2s+2)$, set

$$l(W) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & W & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{k}, \qquad \overbrace{}^{\text{p col.}} \tag{6-132}$$

By (6-129)–(6-132), we have

$$kl(W) = l(W)k. \tag{6-133}$$

Put $w = (\lambda_s, \dots, \lambda_{(p-1)/2}, x) \in \mathbb{C}^{(p-2s+1)/2} \oplus \mathbb{R} \simeq \mathbb{R}^{p-2s+2}$. There exists $W \in \mathfrak{so}(p-2s+2)$ such that

$$\exp(W)w = (0, \dots, 0, |w|), \tag{6-134}$$

where $|w|$ is the Euclidean norm of w.

Put

$$k' = \exp(l(W)) \in K. \tag{6-135}$$

By (6-82), (6-133) and (6-134), we have

$$\mathrm{Ad}(k')a \in \mathfrak{b}, \quad k'kk'^{-1} = k. \tag{6-136}$$

Thus, γ is conjugated by k' into H.

The general case. By (6-63), $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ with $\mathfrak{g}_1 = \mathfrak{sl}_3(\mathbb{R})$ or $\mathfrak{g}_1 = \mathfrak{so}(p,q)$ with $pq > 1$ odd. By (6-98) and (6-121), we have $a \in \mathfrak{g}_1$. The arguments in (6-122)–(6-126) extend directly. We only need to take care of the case $\mathfrak{g}_1 = \mathfrak{so}(p,q)$ and $a^2 = 0$, $v^1 \neq 0$ and $v^2 = 0$. In this case, the arguments in (6-128)–(6-134) extend to the group of isometries G_*. In particular, there is $W_* \in \mathfrak{k}_*$ such that

$$\mathrm{Ad}(\exp(W_*))i_{\mathfrak{g}}(a) \in \mathfrak{b}_*, \quad \mathrm{Ad}(i_G(k))W_* = W_*. \tag{6-137}$$

By (6-94), $\ker(i_{\mathfrak{g}}) \subset \mathfrak{k}$. Let $\ker(i_{\mathfrak{g}})^{\perp}$ be the orthogonal space of $\ker(i_{\mathfrak{g}})$ in $\mathfrak{k}$. Then,

$$\mathfrak{k} = \ker(i_{\mathfrak{g}}) \oplus \ker(i_{\mathfrak{g}})^{\perp}, \quad \ker(i_{\mathfrak{g}})^{\perp} \simeq \mathfrak{k}_*. \tag{6-138}$$

Take $W = (0, W_*) \in \mathfrak{k}$. Put

$$k' = \exp(W) \in K. \tag{6-139}$$

By (6-94), (6-137) and (6-139), we get (6-136). Thus, γ is conjugate by k' into H. The proof of (6-62) is completed. □

7. Selberg and Ruelle zeta functions

In this section, we assume that $\delta(G) = 1$ and that G has compact center. The purpose of this section is to establish the first part of our main result, Theorem 5.5.

In Section 7A, we introduce a class of representations η of M such that $\eta|_{K_M}$ lifts as an element of $\mathrm{RO}(K)$. In particular, η_j is in this class. Take $\hat{\eta} = \Lambda^{\bullet}(\mathfrak{p}_{\mathfrak{m}}^*) \otimes \eta \in \mathrm{RO}(K)$. Using the explicit formulas for orbital integrals of Theorem 4.7, we give an explicit geometric formula for $\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]$, whose proof is given in Section 7B.

In Section 7C, we introduce a Selberg zeta function $Z_{\eta,\rho}$ associated with η and ρ. Using the result in Section 7A, we express $Z_{\eta,\rho}$ in terms of the regularized determinant of the resolvent of $C^{\mathfrak{g},Z,\hat{\eta},\rho}$, and we prove that $Z_{\eta,\rho}$ is meromorphic and satisfies a functional equation.

Finally, in Section 7D, we show that the dynamical zeta function $R_\rho(\sigma)$ is equal to an alternating product of $Z_{\eta_j,\rho}$, from which we deduce the first part of Theorem 5.5.

7A. *An explicit formula for* $\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]$.

We introduce a class of representations of M.

Assumption 7.1. *Let η be a real finite-dimensional representation of M such that*

(1) *the restriction $\eta|_{K_M}$ on K_M can be lifted into* $\mathrm{RO}(K)$;

(2) *the action of the Lie algebra $\mathfrak{u}_{\mathfrak{m}} \subset \mathfrak{m} \otimes_{\mathbb{R}} \mathbb{C}$ on $E_\eta \otimes_{\mathbb{R}} \mathbb{C}$, induced by complexification, can be lifted to an action of Lie group U_M*;

(3) *the Casimir element $C^{\mathfrak{u}_{\mathfrak{m}}}$ of $\mathfrak{u}_{\mathfrak{m}}$ acts on $E_\eta \otimes_{\mathbb{R}} \mathbb{C}$ as the scalar $C^{\mathfrak{u}_{\mathfrak{m}},\eta} \in \mathbb{R}$.*

By Corollary 6.12, let $\hat{\eta} = \hat{\eta}^+ - \hat{\eta}^- \in \mathrm{RO}(K)$ be the virtual real finite-dimensional representation of K on $E_{\hat{\eta}} = E_{\hat{\eta}}^+ - E_{\hat{\eta}}^-$ such that the following identity in $\mathrm{RO}(K_M)$ holds:

$$E_{\hat{\eta}}|_{K_M} = \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} (-1)^i \Lambda^i(\mathfrak{p}_{\mathfrak{m}}^*) \otimes E_\eta|_{K_M}. \tag{7-1}$$

By Corollary 6.12 and by Proposition 6.13, η_j satisfies Assumption 7.1, so that the following identity in $\mathrm{RO}(K)$ holds

$$\sum_{i=1}^{\dim \mathfrak{p}} (-1)^{i-1} i \Lambda^i(\mathfrak{p}^*) = \sum_{j=0}^{2l} (-1)^j E_{\hat{\eta}_i}. \tag{7-2}$$

As in Section 4A, let $\mathcal{E}_{\hat{\eta}} = G \times_K E_{\hat{\eta}}$ be the induced virtual vector bundle on X. Let $C^{\mathfrak{g},X,\hat{\eta}}$ be the corresponding Casimir element of G acting on $C^\infty(X, \mathcal{E}_{\hat{\eta}})$. We will state an explicit formula for $\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]$.

By (6-25), the complex representation of U_M on $E_\eta \otimes_{\mathbb{R}} \mathbb{C}$ extends to a complex representation of $U(\mathfrak{b})$ such that A_0 acts trivially. Set

$$F_{\mathfrak{b},\eta} = U \times_{U(\mathfrak{b})} (E_\eta \otimes_{\mathbb{R}} \mathbb{C}). \tag{7-3}$$

Then $F_{\mathfrak{b},\eta}$ is a complex vector bundle on $Y_{\mathfrak{b}}$. It is equipped with a connection $\nabla^{F_{\mathfrak{b},\eta}}$, induced by $\omega^{\mathfrak{u}(\mathfrak{b})}$, with curvature $R^{F_{\mathfrak{b},\eta}}$.

Remark 7.2. When $\eta = \eta_j$, the above action of $U(\mathfrak{b})$ on $\Lambda^j(\mathfrak{n}_{\mathbb{C}}^*)$ is different from the adjoint action of $U(\mathfrak{b})$ on $\Lambda^j(\mathfrak{n}_{\mathbb{C}}^*)$ induced by (6-31).

Recall that T is the maximal torus of both K and U_M. Put

$$c_G = (-1)^{\frac{m-1}{2}} \frac{|W(T, U_M)|}{|W(T, K)|} \frac{\mathrm{vol}(K/K_M)}{\mathrm{vol}(U_M/K_M)}. \tag{7-4}$$

Recall that $X_M = M/K_M$. By Bott's formula [1965, p. 175],

$$\chi(K/K_M) = \frac{|W(T, K)|}{|W(T, K_M)|}, \tag{7-5}$$

and by (4-5) and (7-4), we have a more geometric expression

$$c_G = (-1)^l \frac{[e(TX_M, \nabla^{TX_M})]^{\max}}{[e(T(K/K_M), \nabla^{T(K/K_M)})]^{\max}}. \tag{7-6}$$

Note that $\dim \mathfrak{u}^\perp(\mathfrak{b}) = 2 \dim \mathfrak{n} = 4l$. If $\beta \in \Lambda^\bullet(\mathfrak{u}^{\perp,*}(\mathfrak{b}))$, let $[\beta]^{\max} \in \mathbb{R}$ be such that

$$\beta - [\beta]^{\max} \frac{\omega^{Y_{\mathfrak{b}},2l}}{(2l)!} \tag{7-7}$$

is of degree smaller than $4l$.

Theorem 7.3. *For $t > 0$, we have*

$$\begin{aligned}\mathrm{Tr_s}^{[1]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] = \frac{c_G}{\sqrt{2\pi t}} \exp\left(\frac{t}{16} \mathrm{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}] - \frac{t}{2} C^{\mathfrak{u}_{\mathfrak{m}},\eta}\right) \\ \times \left[\exp\left(-\frac{\omega^{Y_{\mathfrak{b}},2}}{8\pi^2 |a_0|^2 t}\right) \widehat{A}(TY_{\mathfrak{b}}, \nabla^{TY_{\mathfrak{b}}}) \,\mathrm{ch}(F_{\mathfrak{b},\eta}, \nabla^{F_{\mathfrak{b},\eta}})\right]^{\max}.\end{aligned} \tag{7-8}$$

If $\gamma = e^a k^{-1} \in H$ *with* $a \in \mathfrak{b}$, $a \neq 0$, $k \in T$, *for* $t > 0$, *we have*

$$\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] = \frac{1}{\sqrt{2\pi t}}[e(TX_M(k), \nabla^{TX_M(k)})]^{\max} \times\exp\left(-\frac{|a|^2}{2t}+\frac{t}{16}\mathrm{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}]-\frac{t}{2}C^{\mathfrak{u}_\mathfrak{m},\eta}\right)\frac{\mathrm{Tr}^{E_\eta}[\eta(k^{-1})]}{\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}}}. \tag{7-9}$$

If $\dim \mathfrak{b}(\gamma) \geqslant 2$, *for* $t > 0$, *we have*

$$\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] = 0. \tag{7-10}$$

Proof. The proof of (7-8) and (7-9) will be given in Section 7B. Equation (7-10) is a consequence of (4-22), (6-62) and (7-1). ☐

7B. ***The proof of* (7-8) *and* (7-9).** Let us recall some facts about Lie algebras. Let $\Delta(\mathfrak{t},\mathfrak{k}) \subset \mathfrak{t}^*$ be the real root system [Bröcker and tom Dieck 1985, Definition V.1.3]. We fix a set of positive roots $\Delta^+(\mathfrak{t},\mathfrak{k}) \subset \Delta(\mathfrak{t},\mathfrak{k})$. Set

$$\rho^\mathfrak{k} = \frac{1}{2}\sum_{\alpha\in\Delta^+(\mathfrak{t},\mathfrak{k})}\alpha. \tag{7-11}$$

By Kostant's strange formula [1976] or [Bismut 2011, Proposition 7.5.1], we have

$$4\pi^2|\rho^\mathfrak{k}|^2 = -\tfrac{1}{24}\mathrm{Tr}^\mathfrak{k}[C^{\mathfrak{k},\mathfrak{k}}]. \tag{7-12}$$

Let $\pi_\mathfrak{k} : \mathfrak{t} \to \mathbb{C}$ be the polynomial function such that, for $Y \in \mathfrak{t}$,

$$\pi_\mathfrak{k}(Y) = \prod_{\alpha\in\Delta^+(\mathfrak{t},\mathfrak{k})} 2i\pi\langle\alpha, Y\rangle. \tag{7-13}$$

Let $\sigma_\mathfrak{k} : \mathfrak{t} \to \mathbb{C}$ be the denominator in the Weyl character formula. For $Y \in \mathfrak{t}$, we have

$$\sigma_\mathfrak{k}(Y) = \prod_{\alpha\in\Delta^+(\mathfrak{t},\mathfrak{k})} (e^{i\pi\langle\alpha,Y\rangle} - e^{-i\pi\langle\alpha,Y\rangle}). \tag{7-14}$$

The Weyl group $W(T,K)$ acts isometrically on $\mathfrak{t}$. For $w \in W(T,K)$, set $\epsilon_w = \det(w)|_\mathfrak{t}$. The Weyl denominator formula asserts for $Y \in \mathfrak{t}$, we have

$$\sigma_\mathfrak{k}(Y) = \sum_{w\in W(T,K)} \epsilon_w \exp(2i\pi\langle\rho^\mathfrak{k}, wY\rangle). \tag{7-15}$$

Let $\widehat{K}$ be the set of equivalence classes of complex irreducible representations of K. There is a bijection between $\widehat{K}$ and the set of dominant and analytic integral elements in $\mathfrak{t}^*$ [Bröcker and tom Dieck 1985, Section VI, (1.7)]. If $\lambda \in \mathfrak{t}^*$ is dominant and analytic integral, the character χ_λ of the corresponding complex irreducible representation is given by the Weyl character formula: for $Y \in \mathfrak{t}$,

$$\sigma_\mathfrak{k}(Y)\chi_\lambda(\exp(Y)) = \sum_{w\in W(T,K)} \epsilon_w \exp(2i\pi\langle\rho^\mathfrak{k} + \lambda, wY\rangle). \tag{7-16}$$

Let us recall the Weyl integral formula for Lie algebras. Let $dv_{K/T}$ be the Riemannian volume on K/T induced by $-B$, and let dY be the Lebesgue measure on $\mathfrak{k}$ or $\mathfrak{t}$ induced by $-B$. By [Knapp 1986, Lemma 11.4], if $f \in C_c(\mathfrak{k})$, we have

$$\int_{Y\in\mathfrak{k}} f(Y)\,dY = \frac{1}{|W(T,K)|}\int_{Y\in\mathfrak{t}} |\pi_{\mathfrak{k}}(Y)|^2\left(\int_{k\in K/T} f(\mathrm{Ad}(k)Y)\,dv_{K/T}\right)dY. \tag{7-17}$$

Clearly, the formula (7-17) extends to $L^1(\mathfrak{k})$.

Proof of (7-8). By (3-3), (4-22) and (7-17), we have

$$\begin{aligned}\mathrm{Tr_s}^{[1]}&[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]\\ &= \frac{1}{(2\pi t)^{(m+n)/2}}\exp\left(\frac{t}{16}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right)\\ &\quad\times\frac{\mathrm{vol}(K/T)}{|W(T,K)|}\int_{Y\in\mathfrak{t}} |\pi_{\mathfrak{k}}(Y)|^2 J_1(Y)\,\mathrm{Tr_s}^{E_{\hat{\eta}}}[\exp(-i\hat{\eta}(Y))]\exp(-|Y|^2/(2t))\,dY.\end{aligned} \tag{7-18}$$

As $\delta(M)=0$, we have $\mathfrak{t}$ is also a Cartan subalgebra of $\mathfrak{u}_{\mathfrak{m}}$. We will use (7-17) again to write the integral on the last line of (7-18) as an integral over $\mathfrak{u}_{\mathfrak{m}}$.

By (6-5), we have the isomorphism of representations of K_M,

$$\mathfrak{p}^{\perp}(\mathfrak{b}) \simeq \mathfrak{k}^{\perp}(\mathfrak{b}). \tag{7-19}$$

By (4-21) and (7-19), for $Y\in\mathfrak{t}$, we have

$$J_1(Y) = \frac{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{p}_{\mathfrak{m}}})}{\widehat{A}(i\,\mathrm{ad}(Y)|_{\mathfrak{k}_{\mathfrak{m}}})}. \tag{7-20}$$

By (7-1), for $Y\in\mathfrak{t}$, we have

$$\mathrm{Tr_s}^{E_{\hat{\eta}}}[\exp(-i\hat{\eta}(Y))] = \det(1-\exp(i\,\mathrm{ad}(Y)))|_{\mathfrak{p}_{\mathfrak{m}}}\mathrm{Tr}^{E_{\eta}}[\exp(-i\eta(Y))]. \tag{7-21}$$

By (7-13), (7-20) and (7-21), for $Y\in\mathfrak{t}$, we have

$$\begin{aligned}\frac{|\pi_{\mathfrak{k}}(Y)|^2}{|\pi_{\mathfrak{u}_{\mathfrak{m}}}(Y)|^2}&J_1(Y)\,\mathrm{Tr_s}^{E_{\hat{\eta}}}[\exp(-i\hat{\eta}(Y))]\\ &= (-1)^{\frac{\dim\mathfrak{p}_{\mathfrak{m}}}{2}}\det(\mathrm{ad}(Y))|_{\mathfrak{k}^{\perp}(\mathfrak{b})}\widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_{\mathfrak{m}}})\,\mathrm{Tr}^{E_{\eta}}[\exp(-i\eta(Y))].\end{aligned} \tag{7-22}$$

Using (6-5), for $Y\in\mathfrak{t}$, we have

$$\det(\mathrm{ad}(Y))|_{\mathfrak{k}^{\perp}(\mathfrak{b})} = \det(\mathrm{ad}(Y))|_{\mathfrak{n}_{\mathbb{C}}}. \tag{7-23}$$

By the second condition of Assumption 7.1 and by (7-23), the function on the right-hand side of (7-22) extends naturally to an $\mathrm{Ad}(U_M)$-invariant function defined on $\mathfrak{u}_{\mathfrak{m}}$. By (7-4), (7-17), (7-18), (7-22) and

(7-23), we have

$$
\begin{aligned}
&\mathrm{Tr_s}^{[1]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] \\
&\quad = \frac{(-1)^l c_G}{(2\pi t)^{(m+n)/2}} \exp\left(\frac{t}{16}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right) \\
&\qquad \times \int_{Y\in\mathfrak{u}_\mathfrak{m}} \det(\mathrm{ad}(Y))|_{\mathfrak{n}_\mathbb{C}} \widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_\mathfrak{m}})\,\mathrm{Tr}^{E_\eta}[\exp(-i\eta(Y))]\exp(-|Y|^2/(2t))\,dY.
\end{aligned}
\tag{7-24}
$$

It remains to evaluate the integral on the last line of (7-24). We use the method in [Bismut 2011, Section 7.5]. For $Y \in \mathfrak{u}_\mathfrak{m}$, we have

$$|Y|^2 = -B(Y,Y). \tag{7-25}$$

By (6-32), (6-36) and (6-48), for $Y \in \mathfrak{u}_\mathfrak{m}$, we have

$$B(Y,\Omega^{\mathfrak{u}_\mathfrak{m}}) = -\sum_{1\leqslant i,j\leqslant 2l} B(\mathrm{ad}(Y)f_i,\bar{f}_j)f^i\wedge\bar{f}^j = \sum_{1\leqslant i,j\leqslant 2l}\langle \mathrm{ad}(Y)f_i,f_j\rangle_{\mathfrak{n}_\mathbb{C}} f^i\wedge\bar{f}^j. \tag{7-26}$$

By (6-40), (7-7) and (7-26), for $Y \in \mathfrak{u}_\mathfrak{m}$, we have

$$\frac{\det(\mathrm{ad}(Y))|_{\mathfrak{n}_\mathbb{C}}}{(2\pi t)^{2l}} = (-1)^l\left[\exp\left(\frac{1}{t}B\left(Y,\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}\right)\right)\right]^{\max}. \tag{7-27}$$

As $\dim\mathfrak{u}_\mathfrak{m} = \dim\mathfrak{m} = m+n-2l-1$, from (7-24) and (7-27), we get

$$
\begin{aligned}
&\mathrm{Tr_s}^{[1]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] \\
&\quad = \frac{c_G}{\sqrt{2\pi t}} \exp\left(\frac{t}{16}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right) \\
&\qquad \times \exp\left(\frac{t}{2}\Delta^{\mathfrak{u}_\mathfrak{m}}\right)\left\{\widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_\mathfrak{m}})\,\mathrm{Tr}^{E_\eta}[\exp(-i\eta(Y))]\exp\left(\frac{1}{t}B\left(Y,\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}\right)\right)\right\}^{\max}\Bigg|_{Y=0}.
\end{aligned}
\tag{7-28}
$$

Using

$$B\left(Y,\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}\right) + \frac{1}{2}B(Y,Y) = \frac{1}{2}B\left(Y+\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}, Y+\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}\right) - \frac{1}{2}B\left(\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi},\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}\right), \tag{7-29}$$

by (6-50) and (7-28), we have

$$
\begin{aligned}
&\mathrm{Tr_s}^{[1]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] \\
&\quad = \frac{c_G}{\sqrt{2\pi t}} \exp\left(\frac{t}{16}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right) \\
&\qquad \times\left\{\exp\left(-\frac{\omega^{Y_\mathfrak{b},2}}{8\pi^2|a_0|^2 t}\right)\exp\left(\frac{t}{2}\Delta^{\mathfrak{u}_\mathfrak{m}}\right)(\widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_\mathfrak{m}})\,\mathrm{Tr}^{E_\eta}[\exp(-i\eta(Y))])\right\}^{\max}\Bigg|_{Y=-\frac{\Omega^{\mathfrak{u}_\mathfrak{m}}}{2\pi}}.
\end{aligned}
\tag{7-30}
$$

We claim that the $\mathrm{Ad}(U_M)$-invariant function

$$Y \in \mathfrak{u}_\mathfrak{m} \to \widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_\mathfrak{m}})\mathrm{Tr}^{E_\eta}[\exp(-i\eta(Y))] \tag{7-31}$$

is an eigenfunction of $\Delta^{\mathfrak{u}_\mathfrak{m}}$ with eigenvalue

$$-C^{\mathfrak{u}_\mathfrak{m},\eta}-\tfrac{1}{24}\operatorname{Tr}^{\mathfrak{u}_\mathfrak{m}}[C^{\mathfrak{u}_\mathfrak{m},\mathfrak{u}_\mathfrak{m}}]. \tag{7-32}$$

Indeed, if f is an $\mathrm{Ad}(U_M)$-invariant function on $\mathfrak{u}_\mathfrak{m}$, when restricted to $\mathfrak{t}$, it is well known, for example [Bismut 2011, Equation (7.5.22)], that

$$\Delta^{\mathfrak{u}_\mathfrak{m}} f=\frac{1}{\pi_{\mathfrak{u}_\mathfrak{m}}}\Delta^{\mathfrak{t}}\pi_{\mathfrak{u}_\mathfrak{m}} f. \tag{7-33}$$

Therefore, it is enough to show that the function

$$Y\in\mathfrak{t}\to\pi_{\mathfrak{u}_\mathfrak{m}}(Y)\widehat{A}^{-1}(i\operatorname{ad}(Y))|_{\mathfrak{u}_\mathfrak{m}}\operatorname{Tr}^{E_\eta}[\exp(-i\eta(Y))] \tag{7-34}$$

is an eigenfunction of $\Delta^{\mathfrak{t}}$ with eigenvalue (7-32). For $Y\in\mathfrak{t}$, we have

$$\widehat{A}^{-1}(i\operatorname{ad}(Y))|_{\mathfrak{u}_\mathfrak{m}}=\frac{\sigma_{\mathfrak{u}_\mathfrak{m}}(iY)}{\pi_{\mathfrak{u}_\mathfrak{m}}(iY)}. \tag{7-35}$$

By (7-35), for $Y\in\mathfrak{t}$, we have

$$\pi_{\mathfrak{u}_\mathfrak{m}}(Y)\widehat{A}^{-1}(i\operatorname{ad}(Y))|_{\mathfrak{u}_\mathfrak{m}}=i^{|\Delta^+(\mathfrak{t},\mathfrak{u}_\mathfrak{m})|}\sigma_{\mathfrak{u}_\mathfrak{m}}(-iY). \tag{7-36}$$

If $E_\eta\otimes_\mathbb{R}\mathbb{C}$ is an irreducible representation of U_M with the highest weight $\lambda\in\mathfrak{t}^*$, by the Weyl character formula (7-16), we have

$$\sigma_{\mathfrak{u}_\mathfrak{m}}(-iY)\operatorname{Tr_s}^{E_\eta}[\exp(-i\eta(Y))]=\sum_{w\in W(T,U_M)}\epsilon_w\exp(2\pi\langle\rho^{\mathfrak{u}_\mathfrak{m}}+\lambda,wY\rangle). \tag{7-37}$$

By (7-36) and (7-37), the function (7-34) is an eigenfunction of $\Delta^{\mathfrak{t}}$ with eigenvalue

$$4\pi^2|\rho^{\mathfrak{u}_\mathfrak{m}}+\lambda|^2. \tag{7-38}$$

By Assumption 7.1, the Casimir of $\mathfrak{u}_\mathfrak{m}$ acts as the scalar $C^{\mathfrak{u}_\mathfrak{m},\eta}$. Therefore,

$$-C^{\mathfrak{u}_\mathfrak{m},\eta}=4\pi^2(|\rho^{\mathfrak{u}_\mathfrak{m}}+\lambda|^2-|\rho^{\mathfrak{u}_\mathfrak{m}}|^2). \tag{7-39}$$

By (7-12) and (7-39), the eigenvalue (7-38) is equal to (7-32). If $E_\eta\otimes_\mathbb{R}\mathbb{C}$ is not irreducible, it is enough to decompose $E_\eta\otimes_\mathbb{R}\mathbb{C}$ as a sum of irreducible representations of U_M.

Since the function (7-34) and its derivations of any order satisfy estimations similar to (4-20), by (6-49) and (7-30), we get

$$\begin{aligned}\operatorname{Tr_s}^{[1]}&[\exp(-tC^{\mathfrak{g},X,\hat\eta}/2)]\\ &=\frac{c_G}{\sqrt{2\pi t}}\exp\left(\frac{t}{16}\operatorname{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}]+\frac{t}{48}\operatorname{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]-\frac{t}{48}\operatorname{Tr}^{\mathfrak{u}_\mathfrak{m}}[C^{\mathfrak{u}_\mathfrak{m},\mathfrak{u}_\mathfrak{m}}]-\frac{t}{2}C^{\mathfrak{u}_\mathfrak{m},\eta}\right)\\ &\qquad\times\left\{\exp\left(-\frac{\omega^{Y_\mathfrak{b},2}}{8\pi^2|a_0|^2t}\right)\widehat{A}^{-1}\left(\frac{R^{N_\mathfrak{b}}}{2i\pi}\right)\operatorname{Tr}^{E_\eta}\left[\exp\left(-\frac{R^{F_{\mathfrak{b},\eta}}}{2i\pi}\right)\right]\right\}^{\max}.\end{aligned} \tag{7-40}$$

Since $\widehat{A}$ is an even function, by (2-3), we have

$$\widehat{A}\left(\frac{R^{N_{\mathfrak{b}}}}{2i\pi}\right) = \widehat{A}(N_{\mathfrak{b}}, \nabla^{N_{\mathfrak{b}}}). \tag{7-41}$$

We claim that

$$\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}] - \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{u}_{\mathfrak{m}}}[C^{\mathfrak{u}_{\mathfrak{m}},\mathfrak{u}_{\mathfrak{m}}}] = \mathrm{Tr}^{\mathfrak{u}^{\perp}(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^{\perp}(\mathfrak{b})}]. \tag{7-42}$$

Indeed, by [Bismut 2011, Proposition 2.6.1], we have

$$\begin{aligned}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}] &= \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{u}}[C^{\mathfrak{u},\mathfrak{u}}],\\ \mathrm{Tr}^{\mathfrak{u}^{\perp}(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^{\perp}(\mathfrak{b})}] + \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{u}(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}(\mathfrak{b})}] &= \tfrac{1}{3}\mathrm{Tr}^{\mathfrak{u}}[C^{\mathfrak{u},\mathfrak{u}}].\end{aligned} \tag{7-43}$$

By (6-24), it is trivial that

$$\mathrm{Tr}^{\mathfrak{u}(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}(\mathfrak{b})}] = \mathrm{Tr}^{\mathfrak{u}_{\mathfrak{m}}}[C^{\mathfrak{u}_{\mathfrak{m}},\mathfrak{u}_{\mathfrak{m}}}]. \tag{7-44}$$

From (7-43) and (7-44), we get (7-42).

By (2-4), (6-45) and (7-40)–(7-42), we get (7-8). □

Let $U_M(k)$ be the centralizer of k in U_M, and let $\mathfrak{u}_{\mathfrak{m}}(k)$ be its Lie algebra. Then

$$\mathfrak{u}_{\mathfrak{m}}(k) = \sqrt{-1}\mathfrak{p}_{\mathfrak{m}}(k) \oplus \mathfrak{k}_{\mathfrak{m}}(k). \tag{7-45}$$

Let $U_M^0(k)$ be the connected component of the identity in $U_M(k)$. Clearly, $U_M^0(k)$ is the compact form of $M^0(k)$.

Proof of (7-9). Since $\gamma \in H$, we know $\mathfrak{t} \subset \mathfrak{k}(\gamma)$ is a Cartan subalgebra of $\mathfrak{k}(\gamma)$. By (4-22), (6-17) and (7-17),

$$\begin{aligned}&\mathrm{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]\\ &= \frac{1}{(2\pi t)^{\dim \mathfrak{z}(\gamma)/2}} \exp\left(-\frac{|a|^2}{2t} + \frac{t}{16}\mathrm{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}] + \frac{t}{48}\mathrm{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\right)\\ &\quad\times \frac{\mathrm{vol}(K_M^0(k)/T)}{|W(T,K_M^0(k))|}\int_{Y\in\mathfrak{t}} |\pi_{\mathfrak{k}_{\mathfrak{m}}(k)}(Y)|^2 J_\gamma(Y)\,\mathrm{Tr_s}^{E_{\hat{\eta}}}[\hat{\eta}(k^{-1})\exp(-i\hat{\eta}(Y))]\exp(-|Y|^2/(2t))\,dY.\end{aligned} \tag{7-46}$$

Since $\mathfrak{t}$ is also a Cartan subalgebra of $\mathfrak{u}_{\mathfrak{m}}(k)$, as in the proof of (7-8), we will write the integral on the last line of (7-46) as an integral over $\mathfrak{u}_{\mathfrak{m}}(k)$.

As $k \in T$ and $T \subset K_M$, by (7-1), for $Y \in \mathfrak{t}$, we have

$$\mathrm{Tr_s}^{E_{\hat{\eta}}}[\hat{\eta}(k^{-1})\exp(-i\hat{\eta}(Y))] = \det\big(1-\mathrm{Ad}(k)\exp(i\,\mathrm{ad}(Y))\big)\big|_{\mathfrak{p}_{\mathfrak{m}}} \mathrm{Tr}^{E_\eta}[\eta(k^{-1})\exp(-i\eta(Y))]. \tag{7-47}$$

By (4-19), (7-13) and (7-47), for $Y \in \mathfrak{t}$, we have

$$\begin{aligned}&\frac{|\pi_{\mathfrak{k}_{\mathfrak{m}}(k)}(Y)|^2}{|\pi_{\mathfrak{u}_{\mathfrak{m}}(k)}(Y)|^2} J_\gamma(Y)\,\mathrm{Tr_s}^{E_{\hat{\eta}}}[\hat{\eta}(k^{-1})\exp(-i\hat{\eta}(Y))]\\ &\qquad= \frac{(-1)^{\frac{\dim \mathfrak{p}_{\mathfrak{m}}(k)}{2}}}{\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^{\perp}}\right|^{\frac{1}{2}}}\widehat{A}^{-1}(i\,\mathrm{ad}(Y)|_{\mathfrak{u}_{\mathfrak{m}}(k)})\left[\frac{\det\big(1-\exp(-i\,\mathrm{ad}(Y))\,\mathrm{Ad}(k^{-1})\big)\big|_{\mathfrak{z}_0^{\perp}(\gamma)}}{\det(1-\mathrm{Ad}(k^{-1}))|_{\mathfrak{z}_0^{\perp}(\gamma)}}\right]^{\frac{1}{2}}\\ &\qquad\quad\times \mathrm{Tr}^{E_\eta}[\eta(k^{-1})\exp(-i\eta(Y))].\end{aligned} \tag{7-48}$$

Let $\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)$ be the orthogonal space to $\mathfrak{u}_{\mathfrak{m}}(k)$ in $\mathfrak{u}_{\mathfrak{m}}$. Then

$$\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)=\sqrt{-1}\mathfrak{p}_0^{\perp}(\gamma)\oplus\mathfrak{k}_0^{\perp}(\gamma). \tag{7-49}$$

By (7-49), for $Y\in\mathfrak{t}$, we have

$$\frac{\det\big(1-\exp(-i\operatorname{ad}(Y))\operatorname{Ad}(k^{-1})\big)\big|_{\mathfrak{z}_0^{\perp}(\gamma)}}{\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{z}_0^{\perp}(\gamma)}}=\frac{\det\big(1-\exp(-i\operatorname{ad}(Y))\operatorname{Ad}(k^{-1})\big)\big|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}{\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}. \tag{7-50}$$

By Assumption 7.1 and (7-50), the right-hand side of (7-48) extends naturally to an $\operatorname{Ad}(U_M^0(k))$-invariant function defined on $\mathfrak{u}_{\mathfrak{m}}(k)$. By (4-5), (7-17), (7-46) and (7-48), we have

$$\begin{aligned}
&\operatorname{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)]\\
&=\frac{1}{\sqrt{2\pi t}}\frac{[e(TX_M(k),\nabla^{TX_M(k)})]^{\max}}{\big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{z}_0^{\perp}}\big|^{\frac12}}\exp\Big(-\frac{|a|^2}{2t}+\frac{t}{16}\operatorname{Tr}^{\mathfrak{p}}[C^{\mathfrak{k},\mathfrak{p}}]+\frac{t}{48}\operatorname{Tr}^{\mathfrak{k}}[C^{\mathfrak{k},\mathfrak{k}}]\Big)\exp\Big(\frac{t}{2}\Delta^{\mathfrak{u}_{\mathfrak{m}}(k)}\Big)\\
&\left.\left\{\widehat{A}^{-1}(i\operatorname{ad}(Y)|_{\mathfrak{u}_{\mathfrak{m}}(k)})\left[\frac{\det\big(1-\exp(-i\operatorname{ad}(Y))\operatorname{Ad}(k^{-1})\big)\big|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}{\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}\right]^{\frac12}\operatorname{Tr}^{E_\eta}[\eta(k^{-1})\exp(-i\eta(Y))]\right\}\right|_{Y=0}.
\end{aligned} \tag{7-51}$$

As before, we claim that the function

$$\begin{aligned}
Y\in\mathfrak{u}_{\mathfrak{m}}(k)\to\widehat{A}^{-1}(i\operatorname{ad}(Y)|_{\mathfrak{u}_{\mathfrak{m}}(k)})&\left[\frac{\det\big(1-\exp(-i\operatorname{ad}(Y))\operatorname{Ad}(k^{-1})\big)\big|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}{\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}}\right]^{\frac12}\\
&\times\operatorname{Tr}^{E_\eta}[\eta(k^{-1})\exp(-i\eta(Y))]
\end{aligned} \tag{7-52}$$

is an eigenfunction of $\Delta^{\mathfrak{u}_{\mathfrak{m}}(k)}$ with eigenvalue (7-32). Indeed, it is enough to remark that, as in (7-37), up to a sign, if $k=\exp(\theta_1)$ for some $\theta_1\in\mathfrak{t}$, we have

$$\begin{aligned}
\pi_{\mathfrak{u}_{\mathfrak{m}(k)}}(Y)\widehat{A}^{-1}(i\operatorname{ad}(Y)|_{\mathfrak{u}_{\mathfrak{m}(k)}})\big[\det\big(1-\exp(-i\operatorname{ad}(Y))\operatorname{Ad}(k^{-1})\big)\big|_{\mathfrak{u}_{\mathfrak{m}}^{\perp}(k)}\big]^{\frac12}&\\
=\pm i^{|\Delta^+(\mathfrak{t},\mathfrak{u}_{\mathfrak{m}})|}\sigma_{\mathfrak{u}}(-iY-\theta_1).&
\end{aligned} \tag{7-53}$$

Also, if $E_\eta\otimes_{\mathbb{R}}\mathbb{C}$ is an irreducible representation of U_M with the highest weight $\lambda\in\mathfrak{t}^*$,

$$\sigma_{\mathfrak{u}_{\mathfrak{m}}}(-iY-\theta_1)\operatorname{Tr_s}^{E_\eta}[\eta(k^{-1})\exp(-i\eta(Y))]=\sum_{w\in W(T,U_M)}\epsilon_w\exp\big(2\pi\langle\rho_{\mathfrak{u}_{\mathfrak{m}}}+\lambda,w(Y-i\theta_1)\rangle\big). \tag{7-54}$$

Proceeding as in the proof of (7-8), we get (7-9). □

7C. *Selberg zeta functions.* Recall that $\rho:\Gamma\to\mathrm{U}(r)$ is a unitary representation of Γ and that (F,∇^F,g^F) is the unitarily flat vector bundle on Z associated with ρ.

Definition 7.4. For $\sigma \in \mathbb{C}$, we define a formal sum

$$\Xi_{\eta,\rho}(\sigma) = -\sum_{[\gamma]\in[\Gamma]-\{1\}} \operatorname{Tr}[\rho(\gamma)] \frac{\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})}{m_{[\gamma]}} \frac{\operatorname{Tr}^{E_\eta}[\eta(k^{-1})]}{\left|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}}} e^{-\sigma|a|} \tag{7-55}$$

and a formal Selberg zeta function

$$Z_{\eta,\rho}(\sigma) = \exp(\Xi_{\eta,\rho}(\sigma)). \tag{7-56}$$

The formal Selberg zeta function is said to be well defined if the same conditions as in Definition 5.4 hold.

Remark 7.5. When $G = \mathrm{SO}^0(p,1)$ with $p \geqslant 3$ odd, up to a shift on σ, we know $Z_{\eta,\rho}$ coincides with Selberg zeta function in [Fried 1986, Section 3].

Recall that the Casimir operator $C^{\mathfrak{g},Z,\hat{\eta},\rho}$ acting on $C^\infty(Z, \mathcal{F}_{\hat{\eta}} \otimes_{\mathbb{C}} F)$ is a formally self-adjoint second-order elliptic operator, which is bounded from below. For $\lambda \in \mathbb{C}$, set

$$m_{\eta,\rho}(\lambda) = \dim_{\mathbb{C}} \ker(C^{\mathfrak{g},Z,\hat{\eta}^+,\rho} - \lambda) - \dim_{\mathbb{C}} \ker(C^{\mathfrak{g},Z,\hat{\eta}^-,\rho} - \lambda). \tag{7-57}$$

Write

$$r_{\eta,\rho} = m_{\eta,\rho}(0). \tag{7-58}$$

As in Section 2B, for $\sigma \in \mathbb{R}$ and $\sigma \gg 1$, set

$$\det{}_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta},\rho} + \sigma) = \frac{\det(C^{\mathfrak{g},Z,\hat{\eta}^+,\rho} + \sigma)}{\det(C^{\mathfrak{g},Z,\hat{\eta}^-,\rho} + \sigma)}. \tag{7-59}$$

Then, $\det_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta},\rho} + \sigma)$ extends meromorphically to $\sigma \in \mathbb{C}$. Its zeros and poles belong to the set $\{-\lambda : \lambda \in \operatorname{Sp}(C^{\mathfrak{g},Z,\hat{\eta},\rho})\}$. If $\lambda \in \operatorname{Sp}(C^{\mathfrak{g},Z,\hat{\eta},\rho})$, the order of the zero at $\sigma = -\lambda$ is $m_{\eta,\rho}(\lambda)$.

Set

$$\sigma_\eta = \tfrac{1}{8} \operatorname{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}] - C^{\mathfrak{u}_{\mathfrak{m}},\eta}. \tag{7-60}$$

Set

$$P_\eta(\sigma) = c_G \sum_{j=0}^{l} (-1)^j \frac{\Gamma\left(-j-\frac{1}{2}\right)}{j!(4\pi)^{2j+\frac{1}{2}}|a_0|^{2j}} \left[\omega^{Y_{\mathfrak{b}},2j} \widehat{A}(TY_{\mathfrak{b}}, \nabla^{TY_{\mathfrak{b}}}) \operatorname{ch}(\mathcal{F}_{\mathfrak{b},\eta}, \nabla^{\mathcal{F}_{\mathfrak{b},\eta}})\right]^{\max} \sigma^{2j+1}. \tag{7-61}$$

Then $P_\eta(\sigma)$ is an odd polynomial function of σ. As the notation indicates, σ_η and $P_\eta(\sigma)$ do not depend on Γ or ρ.

Theorem 7.6. *There is $\sigma_0 > 0$ such that*

$$\sum_{[\gamma]\in[\Gamma]-\{1\}} \frac{|\chi_{\mathrm{orb}}(\mathbb{S}^1\backslash B_{[\gamma]})|}{m_{[\gamma]}} \frac{1}{\left|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}}} e^{-\sigma_0|a|} < \infty. \tag{7-62}$$

The Selberg zeta function $Z_{\eta,\rho}(\sigma)$ has a meromorphic extension to $\sigma \in \mathbb{C}$ such that the following identity of meromorphic functions on $\mathbb{C}$ holds:

$$Z_{\eta,\rho}(\sigma) = \det{}_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta},\rho} + \sigma_\eta + \sigma^2) \exp(r \operatorname{vol}(Z) P_\eta(\sigma)). \tag{7-63}$$

The zeros and poles of $Z_{\eta,\rho}(\sigma)$ *belong to the set* $\{\pm i\sqrt{\lambda+\sigma_\eta} : \lambda \in \mathrm{Sp}(C^{\mathfrak{g},Z,\hat{\eta},\rho})\}$. *If* $\lambda \in \mathrm{Sp}(C^{\mathfrak{g},Z,\hat{\eta},\rho})$ *and* $\lambda \neq -\sigma_\eta$, *the order of zero at* $\sigma = \pm i\sqrt{\lambda+\sigma_\eta}$ *is* $m_{\eta,\rho}(\lambda)$. *The order of zero at* $\sigma = 0$ *is* $2m_{\eta,\rho}(-\sigma_\eta)$. *Also*,

$$Z_{\eta,\rho}(\sigma) = Z_{\eta,\rho}(-\sigma)\exp\big(2r\,\mathrm{vol}(Z)P_\eta(\sigma)\big). \tag{7-64}$$

Proof. Proceeding as in the proof of Theorem 5.6, by Proposition 5.1, Corollary 5.2, and Theorem 7.3, we get the first two statements of our theorem. By (7-63), the zeros and poles of $Z_{\eta,\rho}(\sigma)$ coincide with that of $\det_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta},\rho}+\sigma_\eta+\sigma^2)$, from which we deduce the third statement of our theorem. Equation (7-64) is a consequence of (7-63) and of the fact that $P_\eta(\sigma)$ is an odd polynomial. □

7D. ***The Ruelle dynamical zeta function.*** We now consider the Ruelle dynamical zeta function $R_\rho(\sigma)$.

Theorem 7.7. *The dynamical zeta function* $R_\rho(\sigma)$ *is holomorphic for* $\mathrm{Re}(\sigma) \gg 1$ *and extends meromorphically to* $\sigma \in \mathbb{C}$ *such that*

$$R_\rho(\sigma) = \prod_{j=0}^{2l} Z_{\eta_j,\rho}(\sigma+(j-l)|\alpha|)^{(-1)^{j-1}}. \tag{7-65}$$

Proof. Clearly, there is $C > 0$ such that, for all $\gamma \in \Gamma$,

$$\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\right|^{\frac{1}{2}} \leqslant C\exp(C|a|). \tag{7-66}$$

By (7-62) and (7-66), for $\sigma \in \mathbb{C}$ and $\mathrm{Re}(\sigma) > \sigma_0 + C$, the sum in (5-10) converges absolutely to a holomorphic function. By (5-4), (5-7), (5-10), (6-18) and (7-55), for $\sigma \in \mathbb{C}$ and $\mathrm{Re}(\sigma) > \sigma_0 + C$, we have

$$\Xi_\rho(\sigma) = \sum_{j=0}^{2l}(-1)^{j-1}\,\Xi_{\eta_j,\rho}(\sigma+(j-l)|\alpha|). \tag{7-67}$$

By taking exponentials, we get (7-65) for $\mathrm{Re}(\sigma) > \sigma_0 + C$. Since the right-hand side of (7-65) is meromorphic, $R_\rho(\sigma)$ has a meromorphic extension to $\mathbb{C}$ such that (7-65) holds. □

Remark that for $0 \leqslant j \leqslant 2l$, we have the isomorphism of K_M-representations of $\eta_j \simeq \eta_{2l-j}$. By (7-1), we have the isomorphism of K-representations,

$$\hat{\eta}_j \simeq \hat{\eta}_{2l-j}. \tag{7-68}$$

Note that by (6-58) and (7-60), we have

$$\sigma_{\eta_j} = -(j-l)^2|\alpha|^2. \tag{7-69}$$

By (7-63), (7-68) and (7-69), we have

$$\begin{aligned}
&Z_{\eta_j,\rho}\big(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)Z_{\eta_{2l-j},\rho}\big(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)\\
&\qquad= Z_{\eta_j,\rho}\big(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)Z_{\eta_j,\rho}\big(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)\\
&\qquad= \det\nolimits_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta}_j,\rho}+\sigma^2)^2 = \det\nolimits_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta}_j,\rho}+\sigma^2)\det\nolimits_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta}_{2l-j},\rho}+\sigma^2).
\end{aligned} \tag{7-70}$$

Recall that $T(\sigma)$ is defined in (2-14).

Theorem 7.8. *The following identity of meromorphic functions on* $\mathbb{C}$ *holds*:

$$R_\rho(\sigma) = T(\sigma^2)\exp\bigl((-1)^{l-1} r \operatorname{vol}(Z) P_{\eta_l}(\sigma)\bigr) \times \prod_{j=0}^{l-1}\left(\frac{Z_{\eta_j,\rho}(\sigma+(j-l)|\alpha|)Z_{\eta_{2l-j},\rho}(\sigma+(l-j)|\alpha|)}{Z_{\eta_j,\rho}\bigl(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr)Z_{\eta_{2l-j},\rho}\bigl(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr)}\right)^{(-1)^{j-1}}. \tag{7-71}$$

Proof. By (2-14), (4-24), and (7-59), we have the identity of meromorphic functions,

$$T(\sigma) = \prod_{j=0}^{2l} \det\nolimits_{\mathrm{gr}}(C^{\mathfrak{g},Z,\hat{\eta}_j,\rho}+\sigma)^{(-1)^{j-1}}. \tag{7-72}$$

By (7-63), (7-70), and (7-72), we have

$$T(\sigma^2) = Z_{\eta_l,\rho}(\sigma)^{(-1)^{l-1}} \exp\bigl((-1)^l r \operatorname{vol}(Z) P_{\eta_l}(\sigma)\bigr) \times \prod_{j=0}^{l-1}\bigl(Z_{\eta_j,\rho}\bigl(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr)Z_{\eta_{2l-j},\rho}\bigl(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr)\bigr)^{(-1)^{j-1}}. \tag{7-73}$$

By (7-65) and (7-73), we get (7-71). □

For $0 \leqslant j \leqslant 2l$, as in (7-58), we write $r_j = r_{\eta_j,\rho}$. By (7-68) and (7-72), we have

$$\chi'(Z,F) = 2\sum_{j=0}^{l-1}(-1)^{j-1} r_j + (-1)^{l-1} r_l. \tag{7-74}$$

Set

$$C_\rho = \prod_{j=0}^{l-1}(-4(l-j)^2|\alpha|^2)^{(-1)^{j-1}r_j}, \quad r_\rho = 2\sum_{j=0}^{l}(-1)^{j-1} r_j. \tag{7-75}$$

Proof of (5-12). By Proposition 6.13 and Theorem 7.6, for $0 \leqslant j \leqslant l-1$, the orders of the zeros at $\sigma = 0$ of the functions $Z_{\eta_j,\rho}(\sigma+(j-l)|\alpha|)$ and $Z_{\eta_{2l-j},\rho}(\sigma+(l-j)|\alpha|)$ are equal to r_j. Therefore, for $0 \leqslant j \leqslant l-1$, there are $A_j \neq 0$, $B_j \neq 0$ such that, as $\sigma \to 0$,

$$\begin{aligned} Z_{\eta_j,\rho}(\sigma+(j-l)|\alpha|) &= A_j\sigma^{r_j} + \mathcal{O}(\sigma^{r_j+1}),\\ Z_{\eta_{2l-j},\rho}(\sigma+(l-j)|\alpha|) &= B_j\sigma^{r_j} + \mathcal{O}(\sigma^{r_j+1}), \end{aligned} \tag{7-76}$$

and

$$\begin{aligned} Z_{\eta_j,\rho}\bigl(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr) &= A_j\left(\frac{-\sigma^2}{2(l-j)|\alpha|}\right)^{r_j} + \mathcal{O}(\sigma^{2r_j+2}),\\ Z_{\eta_{2l-j},\rho}\bigl(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\bigr) &= B_j\left(\frac{\sigma^2}{2(l-j)|\alpha|}\right)^{r_j} + \mathcal{O}(\sigma^{2r_j+2}). \end{aligned} \tag{7-77}$$

By (7-76) and (7-77), as $\sigma \to 0$,

$$\frac{Z_{\eta_j,\rho}(\sigma+(j-l)|\alpha|)Z_{\eta_{2l-j},\rho}(\sigma+(l-j)|\alpha|)}{Z_{\eta_j,\rho}\big(-\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)Z_{\eta_{2l-j},\rho}\big(\sqrt{\sigma^2+(l-j)^2|\alpha|^2}\big)} \to (-4(l-j)^2|\alpha|^2)^{r_j}\sigma^{-2r_j}+\mathcal{O}(\sigma^{-2r_j+1}). \tag{7-78}$$

By (7-61), (7-71), (7-74), (7-75) and (7-78), we get (5-12). □

Remark 7.9. When $G = \mathrm{SO}^0(p,1)$ with $p \geqslant 3$ odd, we recover [Fried 1986, Theorem 3].

Remark 7.10. If we scale the form B with the factor $a > 0$, then $R_\rho(\sigma)$ is replaced by $R_\rho(\sqrt{a}\sigma)$. By (5-12), as $\sigma \to 0$,

$$R_\rho(\sqrt{a}\sigma) = a^{\frac{r_\rho}{2}} C_\rho T(F)^2 \sigma^{r_\rho} + \mathcal{O}(\sigma^{r_\rho+1}). \tag{7-79}$$

On the other hand, C_ρ should become $a^{\sum_{j=0}^{l-1}(-1)^j r_j} C_\rho$, and $T(F)$ should scale by $a^{\chi'(Z,F)/2}$. This is only possible if

$$r_\rho = 2\sum_{j=0}^{l-1}(-1)^j r_j + 2\chi'(Z,F), \tag{7-80}$$

which is just (7-74).

8. A cohomological formula for r_j

The purpose of this section is to establish (5-13) when G has compact center and is such that $\delta(G) = 1$. We rely on some deep results from the representation theory of reductive Lie groups.

This section is organized as follows. In Section 8A, we recall the constructions of the infinitesimal and global characters of Harish-Chandra modules. We also recall some properties of $(\mathfrak{g}, K)$-cohomology and $\mathfrak{n}$-homology of Harish-Chandra modules.

In Section 8B, we give a formula relating r_j with an alternating sum of the dimensions of Lie algebra cohomologies of certain Harish-Chandra modules, and we establish (5-13).

8A. ***Some results from representation theory.*** In this section, we do not assume that $\delta(G) = 1$. We use the notation in Section 3 and the convention of real root systems introduced in Section 7B.

8A1. *Infinitesimal characters.* Let $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$ be the center of the enveloping algebra $U(\mathfrak{g}_\mathbb{C})$ of the complexification $\mathfrak{g}_\mathbb{C}$ of $\mathfrak{g}$. A morphism of algebras $\chi : \mathcal{Z}(\mathfrak{g}_\mathbb{C}) \to \mathbb{C}$ will be called a character of $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$.

Recall that $\mathfrak{h}_1, \dots, \mathfrak{h}_{l_0}$ form all the nonconjugated θ-stable Cartan subalgebras of $\mathfrak{g}$. Let $\mathfrak{h}_{i\mathbb{C}} = \mathfrak{h}_i \otimes_\mathbb{R} \mathbb{C}$ and $\mathfrak{h}_{i\mathbb{R}} = \sqrt{-1}\mathfrak{h}_{i\mathfrak{p}} \oplus \mathfrak{h}_{i\mathfrak{k}}$ be the complexification and real form of $\mathfrak{h}_i$. For $\alpha \in \mathfrak{h}_{i\mathbb{R}}^*$, we extend α to a $\mathbb{C}$-linear form on $\mathfrak{h}_{i\mathbb{C}}$ by $\mathbb{C}$-linearity. In this way, we identify $\mathfrak{h}_{i\mathbb{R}}^*$ to a subset of $\mathfrak{h}_{i\mathbb{C}}^*$.

For $1 \leqslant i \leqslant l_0$, let $S(\mathfrak{h}_{i\mathbb{C}})$ be the symmetric algebra of $\mathfrak{h}_{i\mathbb{C}}$. The algebraic Weyl group $W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$ acts isometrically on $\mathfrak{h}_{i\mathbb{R}}$. By $\mathbb{C}$-linearity, $W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$ acts on $\mathfrak{h}_{i\mathbb{C}}$. Therefore, $W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$ acts on $S(\mathfrak{h}_{i\mathbb{C}})$. Let $S(\mathfrak{h}_{i\mathbb{C}})^{W(\mathfrak{h}_{i\mathbb{R}},\mathfrak{u})} \subset S(\mathfrak{h}_{i\mathbb{C}})$ be the $W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$-invariant subalgebra of $S(\mathfrak{h}_{i\mathbb{C}})$. Let

$$\gamma_i : \mathcal{Z}(\mathfrak{g}_\mathbb{C}) \simeq S(\mathfrak{h}_{i\mathbb{C}})^{W(\mathfrak{h}_{i\mathbb{R}},\mathfrak{u})} \tag{8-1}$$

be the Harish-Chandra isomorphism [Knapp 2002, Section V.5]. For $\Lambda \in \mathfrak{h}_{i\mathbb{C}}^*$, we can associate to it a character χ_Λ of $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$ as follows: for $z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$,

$$\chi_\Lambda(z) = \langle \gamma_i(z), 2\sqrt{-1}\pi\Lambda\rangle. \tag{8-2}$$

By [Knapp 2002, Theorem 5.62], every character of $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$ is of the form χ_Λ for some $\Lambda \in \mathfrak{h}_{i\mathbb{C}}^*$. Also, Λ is uniquely determined up to an action of $W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$. Such an element $\Lambda \in \mathfrak{h}_{i\mathbb{C}}^*$ is called the Harish-Chandra parameter of the character. In particular, $\chi_\Lambda = 0$ if and only if there is $w \in W(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$ such that

$$w\Lambda = \rho_i^\mathfrak{u}, \tag{8-3}$$

where $\rho_i^\mathfrak{u}$ is defined as in (7-11) with respect to $(\mathfrak{h}_{i\mathbb{R}}, \mathfrak{u})$.

Definition 8.1. A complex representation of $\mathfrak{g}_\mathbb{C}$ is said to have infinitesimal character χ if $z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$ acts as a scalar $\chi(z) \in \mathbb{C}$.

A complex representation of $\mathfrak{g}_\mathbb{C}$ is said to have generalized infinitesimal character χ if $z - \chi(z)$ acts nilpotently for all $z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$, i.e., $(z - \chi(z))^i$ acts like 0 for $i \gg 1$.

If $\lambda \in \mathfrak{h}_{i\mathbb{R}}^*$ is algebraically integral and dominant, let V_λ be the complex finite-dimensional irreducible representation of the Lie algebra $\mathfrak{g}_\mathbb{C}$ with the highest weight λ. Then V_λ possesses an infinitesimal character with Harish-Chandra parameter $\lambda + \rho_i^\mathfrak{u} \in \mathfrak{h}_{i\mathbb{R}}^*$.

8A2. *Harish-Chandra* $(\mathfrak{g}_\mathbb{C}, K)$*-modules and admissible representations of* G. We follow [Hecht and Schmid 1983, pp. 54–55] and [Knapp 1986, p. 207].

Definition 8.2. We will say that a complex $U(\mathfrak{g}_\mathbb{C})$-module V, equipped with an action of K, is a Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-module, if the following conditions hold:

(1) The space V is finitely generated as a $U(\mathfrak{g}_\mathbb{C})$-module.

(2) Every $v \in V$ lies in a finite-dimensional, $\mathfrak{k}_\mathbb{C}$-invariant subspace.

(3) The actions of $\mathfrak{g}_\mathbb{C}$ and K are compatible.

(4) Each irreducible K-module occurs only finitely many times in V.

Let V be a Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-module. For a character χ of $\mathcal{Z}(\mathfrak{g}_\mathbb{C})$, let $V_\chi \subset V$ be the largest submodule of V on which $z - \chi(z)$ acts nilpotently for all $z \in \mathcal{Z}(\mathfrak{g}_\mathbb{C})$. Then V_χ is a Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-submodule of V with generalized infinitesimal character χ. By [Hecht and Schmid 1983, Equation (2.4)], we can decompose V as a finite sum of Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-submodules

$$V = \bigoplus_\chi V_\chi. \tag{8-4}$$

Any Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-module V has a finite composition series in the following sense: there exist finitely many Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-submodules

$$V = V_{n_1} \supset V_{n_1-1} \supset \cdots \supset V_0 \supset V_{-1} = 0 \tag{8-5}$$

such that each quotient V_i/V_{i-1}, for $0 \leqslant i \leqslant n_1$, is an irreducible Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module. Moreover, the set of all irreducible quotients and their multiplicities are the same for all the composition series.

Definition 8.3. We say that a representation π of G on a Hilbert space is admissible if the following hold:

(1) When restricted to K, $\pi|_K$ is unitary.

(2) Each $\tau \in \widehat{K}$ occurs with only finite multiplicity in $\pi|_K$.

Let π be a finitely generated admissible representation of G on the Hilbert space V_π. If $\tau \in \widehat{K}$, let $V_\pi(\tau) \subset V_\pi$ be the τ-isotopic subspace of V_π. Then $V_\pi(\tau)$ is the image of the evaluation map

$$(f, v) \in \mathrm{Hom}_K(V_\tau, V_\pi) \otimes V_\tau \to f(v) \in V_\pi. \tag{8-6}$$

Let

$$V_{\pi,K} = \bigoplus_{\tau \in \widehat{K}} V_\pi(\tau) \subset V_\pi \tag{8-7}$$

be the algebraic sum of representations of K. By [Knapp 1986, Proposition 8.5], $V_{\pi,K}$ is a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module. It is explained in [Vogan 2008, Section 4] that, by results of Casselman, Harish-Chandra, Lepowsky and Wallach, any Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module V can be constructed in this way and the corresponding V_π is called a Hilbert globalization of V. Moreover, V is an irreducible Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module if and only if V_π is an irreducible admissible representation of G. In this case, V or V_π has an infinitesimal character.

We note that a Hilbert globalization of a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module is not unique.

8A3. *Global characters.* We recall the definition of the space of rapidly decreasing functions $\mathcal{S}(G)$ on G [Wallach 1988, Section 7.1.2].

For $z \in U(\mathfrak{g})$, we denote by z_L and z_R respectively the corresponding left and right invariant differential operators on G. For $r \geqslant 0$, $z_1 \in U(\mathfrak{g})$, $z_2 \in U(\mathfrak{g})$ and $f \in C^\infty(G)$, put

$$\|f\|_{r,z_1,z_2} = \sup_{g \in G} e^{r d_X(p1, pg)} |z_{1L} z_{2R} f(g)|. \tag{8-8}$$

Let $\mathcal{S}(G)$ be the space of all $f \in C^\infty(G)$ such that, for all $r \geqslant 0$, $z_1 \in U(\mathfrak{g})$, $z_2 \in U(\mathfrak{g})$, we have $\|f\|_{r,z_1,z_2} < \infty$. We endow $\mathcal{S}(G)$ with the topology given by the above seminorms. By [Wallach 1988, Theorem 7.1.1], $\mathcal{S}(G)$ is a Fréchet space which contains $C_c^\infty(G)$ as a dense subspace.

Let π be a finitely generated admissible representation of G on the Hilbert space V_π. By [Wallach 1988, Lemma 2.A.2.2], there exists $C > 0$ such that, for $g \in G$, we have

$$\|\pi(g)\| \leqslant C e^{C d_X(p1, pg)}, \tag{8-9}$$

where $\|\cdot\|$ is the operator norm. By (8-9), if $f \in \mathcal{S}(G)$,

$$\pi(f) = \int_G f(g)\pi(g)\, dg \tag{8-10}$$

is a bounded operator on V_π. By [Wallach 1988, Lemma 8.1.1], $\pi(f)$ is trace class. The global character Θ_π^G of π is a continuous linear functional on $\mathcal{S}(G)$ such that, for $f \in \mathcal{S}(G)$,

$$\operatorname{Tr}[\pi(f)] = \langle \Theta_\pi^G, f\rangle. \tag{8-11}$$

If V is a Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-module, we can define the global character Θ_V^G of V by the global character of its Hilbert globalization. We note that the global character does not depend on the choice of Hilbert globalization [Hecht and Schmid 1983, p. 56].

By Harish-Chandra's regularity theorem [Knapp 1986, Theorems 10.25], there is an L^1_{loc} and $\mathrm{Ad}(G)$-invariant function $\Theta_\pi^G(g)$ on G, whose restriction to the regular set G' is analytic, such that, for $f \in C_c^\infty(G)$, we have

$$\langle \Theta_\pi^G, f\rangle = \int_{g\in G} \Theta_\pi^G(g) f(g)\, dv_G. \tag{8-12}$$

Proposition 8.4. *If $f \in \mathcal{S}(G)$, then $\Theta_\pi^G(g) f(g) \in L^1(G)$ such that*

$$\langle \Theta_\pi^G, f\rangle = \int_{g\in G} \Theta_\pi^G(g) f(g)\, dv_G. \tag{8-13}$$

Proof. It is enough to show that there exist $C > 0$ and a seminorm $\|\cdot\|$ on $\mathcal{S}(G)$ such that

$$\int_G |\Theta_\pi^G(g) f(g)|\, dg \leqslant C\|f\|. \tag{8-14}$$

Recall that H' is defined in (3-36). By (3-33), we need to show that there exist $C > 0$ and a seminorm $\|\cdot\|$ on $\mathcal{S}(G)$ such that, for $1 \leqslant i \leqslant l_0$, we have

$$\int_{\gamma\in H_i'} |\Theta_\pi^G(\gamma)| \left(\int_{g\in H_i\backslash G} |f(g^{-1}\gamma g)|\, dv_{H_i\backslash G}\right) \left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\right| dv_{H_i} \leqslant C\|f\|. \tag{8-15}$$

By [Knapp 1986, Theorem 10.35], there exist $C > 0$ and $r_0 > 0$ such that, for $\gamma = e^a k^{-1} \in H_i'$ with $a \in \mathfrak{h}_{i\mathfrak{p}}$, $k \in H_i \cap K$, we have

$$|\Theta_\pi^G(\gamma)| \left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\right|^{\frac{1}{2}} \leqslant C e^{r_0|a|}. \tag{8-16}$$

We claim that there exist $r_1 > 0$ and $C > 0$, such that, for $\gamma \in H_i'$, we have

$$\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\right|^{\frac{1}{2}} \int_{g\in H_i\backslash G} \exp\bigl(-r_1 d_X(p1, g^{-1}\gamma g\cdot p1)\bigr)\, dv_{H_i\backslash G} \leqslant C. \tag{8-17}$$

Indeed, let $\Xi(g)$ be the Harish-Chandra Ξ-function [Varadarajan 1977, Section II.8.5]. By Section II.12.2 and Corollary 5 of the same paper, there exist $r_2 > 0$ and $C > 0$, such that, for $\gamma \in H_i'$, we have

$$\left|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\right|^{\frac{1}{2}} \int_{g\in H_i\backslash G} \Xi(g^{-1}\gamma g)\bigl(1 + d_X(p1, g^{-1}\gamma g\cdot p1)\bigr)^{-r_2}\, dv_{H_i\backslash G} \leqslant C. \tag{8-18}$$

By [Knapp 1986, Proposition 7.15(c)] and by (8-18), we get (8-17).

By [Bismut 2011, Equation (3.1.10)], for $g \in G$ and $\gamma = e^a k^{-1} \in H_i$, we have

$$d_X(p1, g^{-1}\gamma g\cdot p1) \geqslant |a|. \tag{8-19}$$

Take $r=2r_0+r_1$, $z_1=z_2=1\in U(\mathfrak{g})$. Since $f\in\mathcal{S}(G)$, by (8-8) and (8-19), for $\gamma=e^ak^{-1}\in H_i'$, we have

$$\begin{aligned}|f(g^{-1}\gamma g)| &\leqslant \|f\|_{r,z_1,z_2}\exp\bigl(-rd_X(p1,g^{-1}\gamma g\cdot p1)\bigr)\\ &\leqslant \|f\|_{r,z_1,z_2}\exp(-2r_0|a|)\exp\bigl(-r_1d_X(p1,g^{-1}\gamma g\cdot p1)\bigr).\end{aligned}\tag{8-20}$$

By (8-16), (8-17) and (8-20), for $\gamma\in H_i'$, we have

$$|\Theta_\pi^G(\gamma)|\biggl(\int_{g\in H_i\backslash G}|f(g^{-1}\gamma g)|\,dv_{H_i\backslash G}\biggr)\bigl|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\bigr|\leqslant C\|f\|_{r,z_1,z_2}\exp(-r_0|a|).\tag{8-21}$$

By (8-21), we get (8-15). □

Let V be a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}},K)$-module, and let τ be a real finite-dimensional orthogonal representation of K on the real Euclidean space E_τ. Then the invariant subspace $(V\otimes_{\mathbb{R}}E_\tau)^K\subset V\otimes_{\mathbb{R}}E_\tau$ has a finite dimension. We will describe an integral formula for $\dim_{\mathbb{C}}(V\otimes_{\mathbb{R}}E_\tau)^K$, which extends [Barbasch and Moscovici 1983, Corollary 2.2].

Recall that $p_t^{X,\tau}(g)$ is the smooth integral kernel of $\exp(-tC^{X,\tau}/2)$. By the estimation on the heat kernel or by [Barbasch and Moscovici 1983, Proposition 2.4], $p_t^{X,\tau}(g)\in\mathcal{S}(G)\otimes\mathrm{End}(E_\tau)$. Recall that dv_G is the Riemannian volume on G induced by $-B(\cdot,\theta\cdot)$.

Proposition 8.5. *Let $f\in C^\infty(G,E_\tau)^K$. Assume that there exist $C>0$ and $r>0$ such that*

$$|f(g)|\leqslant C\exp(rd_X(p1,pg)).\tag{8-22}$$

The integral

$$\int_{g\in G}p_t^{X,\tau}(g)f(g)\,dv_G\in E_\tau\tag{8-23}$$

is well defined so that

$$\begin{aligned}\frac{\partial}{\partial t}\int_{g\in G}p_t^{X,\tau}(g)f(g)\,dv_G&=-\frac{1}{2}\int_{g\in G}C^{\mathfrak{g}}p_t^{X,\tau}(g)f(g)\,dv_G,\\ \frac{1}{\mathrm{vol}(K)}\lim_{t\to0}\int_{g\in G}p_t^{X,\tau}(g)f(g)\,dv_G&=f(1).\end{aligned}\tag{8-24}$$

Proof. By (8-22), by the property of $\mathcal{S}(G)$ and by $\frac{\partial}{\partial t}p_t^{X,\tau}(g)=-\frac{1}{2}C^{\mathfrak{g}}p_t^{X,\tau}(g)$, the left-hand side of (8-23) and the right-hand side of the first equation of (8-24) are well defined so that the first equation of (8-24) holds true.

It remains to show the second equation of (8-24). Let $\phi_1\in C_c^\infty(G)^K$ be such that $0\leqslant\phi_1(g)\leqslant1$ and

$$\phi_1(g)=\begin{cases}1, & d_X(p1,pg)\leqslant1,\\ 0, & d_X(p1,pg)\geqslant2.\end{cases}\tag{8-25}$$

Set $\phi_2=1-\phi_1$.

Since ϕ_1f has compact support, it descends to an L^2-section on X with values in $G\times_KE_\tau$. We have

$$\frac{1}{\mathrm{vol}(K)}\lim_{t\to0}\int_{g\in G}p_t^{X,\tau}(g)\phi_1(g)f(g)\,dv_G=f(1).\tag{8-26}$$

By (4-27), there exist $c > 0$ and $C > 0$ such that, for $g \in G$ with $d_X(p1, pg) \geqslant 1$ and for $t \in (0, 1]$, we have

$$|p_t^{X,\tau}(g)| \leqslant C \exp\left(-c \frac{d_X^2(p1, pg)}{t}\right) \leqslant C e^{-\frac{c}{2t}} \exp\left(-c \frac{d_X^2(p1, pg)}{2t}\right). \tag{8-27}$$

By (8-22) and (8-27), there exist $c > 0$ and $C > 0$ such that, for $t \in (0, 1]$, we have

$$\int_{g \in G} \left| p_t^{X,\tau}(g) \phi_2(g) f(g)\, dv_G \right| \leqslant C e^{-\frac{c}{2t}}. \tag{8-28}$$

By (8-26) and (8-28), we get the second equation of (8-24). □

Proposition 8.6. *Let V be a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module with generalized infinitesimal character χ. For $t > 0$, we have*

$$\dim_{\mathbb{C}} (V \otimes_{\mathbb{R}} E_\tau)^K = \mathrm{vol}(K)^{-1} e^{\frac{t\chi(C^{\mathfrak{g}})}{2}} \int_{g \in G} \Theta_V^G(g) \operatorname{Tr}[p_t^{X,\tau}(g)]\, dv_G. \tag{8-29}$$

Proof. Let V_π be a Hilbert globalization of V. Then,

$$(V \otimes_{\mathbb{R}} E_\tau)^K = (V_\pi \otimes_{\mathbb{R}} E_\tau)^K. \tag{8-30}$$

As in (8-10), set

$$\pi(p_t^{X,\tau}) = \frac{1}{\mathrm{vol}(K)} \int_{g \in G} \pi(g) \otimes_{\mathbb{R}} p_t^{X,\tau}(g)\, dv_G. \tag{8-31}$$

Then, $\pi(p_t^{X,\tau})$ is a bounded operator acting on $V_\pi \otimes_{\mathbb{R}} E_\tau$.

We follow [Barbasch and Moscovici 1983, pp. 160–161]. Let $(V_\pi \otimes_{\mathbb{R}} E_\tau)^{K,\perp}$ be the orthogonal space to $(V_\pi \otimes_{\mathbb{R}} E_\tau)^K$ in $V_\pi \otimes_{\mathbb{R}} E_\tau$ such that

$$V_\pi \otimes_{\mathbb{R}} E_\tau = (V_\pi \otimes_{\mathbb{R}} E_\tau)^K \oplus (V_\pi \otimes_{\mathbb{R}} E_\tau)^{K,\perp}. \tag{8-32}$$

Let $Q_{\pi,\tau}$ be the orthogonal projection from $V_\pi \otimes_{\mathbb{R}} E_\tau$ to $(V_\pi \otimes_{\mathbb{R}} E_\tau)^K$. Then,

$$Q_{\pi,\tau} = \frac{1}{\mathrm{vol}(K)} \int_{k \in K} \pi \otimes \tau(k)\, dv_K. \tag{8-33}$$

By (4-13), (8-31) and (8-33), we get

$$Q_{\pi,\tau} \pi(p_t^{X,\tau}) Q_{\pi,\tau} = \pi(p_t^{X,\tau}). \tag{8-34}$$

In particular, $\pi(p_t^{X,\tau})$ is of finite rank.

Take $u \in (V_\pi \otimes_{\mathbb{R}} E_\tau)^K$ and $v \in V_\pi$. Define $\langle u, v \rangle \in E_\tau$ to be such that, for any $w \in E_\tau$,

$$\langle \langle u, v \rangle, w \rangle = \langle u, v \otimes_{\mathbb{R}} w \rangle. \tag{8-35}$$

By (8-9), the function $g \in G \to \langle \pi(g) \otimes_{\mathbb{R}} \mathrm{id} \cdot u, v \rangle \in E_\tau$ is of class $C^\infty(G, E_\tau)^K$ such that (8-22) holds.

By (8-31), we have

$$\langle \pi(p_t^{X,\tau})u, v\rangle = \frac{1}{\mathrm{vol}(K)} \int_{g\in G} p_t^{X,\tau}(g)\big\langle \pi(g)\otimes_{\mathbb{R}} \mathrm{id}\cdot u, v\big\rangle \, dv_G. \tag{8-36}$$

By Proposition 8.5 and (8-36), we have

$$\frac{\partial}{\partial t}\langle \pi(p_t^{X,\tau})u, v\rangle = -\frac{1}{2}\big\langle \pi(C^{\mathfrak{g}})\pi(p_t^{X,\tau})u, v\big\rangle, \quad \lim_{t\to 0}\langle \pi(p_t^{X,\tau})u, v\rangle = \langle u, v\rangle. \tag{8-37}$$

Since $C^{\mathfrak{g}} \in \mathcal{Z}(\mathfrak{g})$ and since $\pi(C^{\mathfrak{g}})$ preserves the splitting (8-32), by (8-34) and (8-37), under the splitting (8-32), we have

$$\pi(p_t^{X,\tau}) = \begin{pmatrix} e^{-t\pi(C^{\mathfrak{g}})/2} & 0 \\ 0 & 0 \end{pmatrix}. \tag{8-38}$$

Since V has a generalized infinitesimal character χ, by (8-38), we have

$$\mathrm{Tr}[\pi(p_t^{X,\tau})] = e^{-t\chi(C^{\mathfrak{g}})/2} \dim_{\mathbb{C}}(V_\pi \otimes_{\mathbb{R}} E_\tau)^K. \tag{8-39}$$

Let $(\xi_i)_{i=1}^{\infty}$ and $(\eta_j)_{j=1}^{\dim E_\tau}$ be orthogonal bases of V_π and E_τ. Then

$$\mathrm{Tr}[\pi(p_t^{X,\tau})] = \frac{1}{\mathrm{vol}(K)} \sum_{i=1}^{\infty} \sum_{j=1}^{\dim E_\tau} \int_{g\in G} \langle p_t^{X,\tau}(g)\eta_j, \eta_j\rangle \langle \pi(g)\xi_i, \xi_i\rangle \, dv_G \tag{8-40}$$

$$= \frac{1}{\mathrm{vol}(K)} \sum_{i=1}^{\infty} \int_{g\in G} \mathrm{Tr}[p_t^{X,\tau}(g)]\langle \pi(g)\xi_i, \xi_i\rangle \, dv_G. \tag{8-41}$$

Since $\mathrm{Tr}[p_t^{X,\tau}(g)] \in \mathcal{S}(G)$, by (8-13) and (8-40), we have

$$\mathrm{Tr}[\pi(p_t^{X,\tau})] = \frac{1}{\mathrm{vol}(K)} \int_{g\in G} \mathrm{Tr}[p_t^{X,\tau}(g)]\Theta_\pi^G(g) \, dv_G. \tag{8-42}$$

From (8-30), (8-39) and (8-42), we get (8-29). □

Proposition 8.7. *For $1 \leqslant i \leqslant l_0$, the function*

$$\gamma \in H_i' \to \mathrm{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]\Theta_\pi^G(\gamma)\big|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\big| \tag{8-43}$$

is almost everywhere well defined and integrable on H_i' so that

$$\int_{g\in G} \mathrm{Tr}[p_t^{X,\tau}(g)]\Theta_\pi^G(g)\, dv_G$$
$$= \sum_{i=1}^{l_0} \frac{\mathrm{vol}(K\cap H_i\backslash K)}{|W(H_i, G)|} \int_{\gamma\in H_i'} \mathrm{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]\Theta_\pi^G(g)\big|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\big| \, dv_{H_i}. \tag{8-44}$$

Proof. Since $\mathrm{Tr}[p_t^{X,\tau}(g)]\Theta_\pi^G(g) \in L^1(G)$, by (3-33) and by Fubini's theorem, the function

$$\gamma \in H_i \to \left(\int_{g\in H_i\backslash G} \mathrm{Tr}^{E_\tau}[p_t^{X,\tau}(g^{-1}\gamma g)]\, dv_{H_i\backslash G}\right)\Theta_\pi^G(\gamma)\big|\det(1-\mathrm{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}_i}\big| \tag{8-45}$$

is almost everywhere well defined and integrable on H_i.

Take $\gamma \in H_i'$. Since H_i is abelian, we have

$$Z^0(\gamma) = H_i^0 \subset H_i \subset Z(\gamma). \tag{8-46}$$

We have a finite covering space $H_i^0\backslash G \to H_i\backslash G$. Note that

$$[H_i : H_i^0] = [K \cap H_i : K \cap H_i^0]. \tag{8-47}$$

By (4-15), (8-46) and (8-47), if $\gamma \in H_i'$, we have

$$\begin{aligned}\int_{H_i\backslash G} \operatorname{Tr}^{E_\tau}[p_t^{X,\tau}(g^{-1}\gamma g)]\, dv_{H_i\backslash G} &= \frac{\operatorname{vol}(K^0(\gamma)\backslash K)}{[H_i : H_i^0]} \operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)] \\ &= \operatorname{vol}(K \cap H_i\backslash K) \operatorname{Tr}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\tau}/2)]. \end{aligned} \tag{8-48}$$

Since $H_i - H_i'$ has zero measure, and by (8-45) and (8-48), the function (8-43) defines an L^1-function on H_i'. By (3-33) and (8-48), we get (8-44). □

8A4. *The* $(\mathfrak{g}, K)$*-cohomology.* If V is a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module, let $H^\bullet(\mathfrak{g}, K; V)$ be the $(\mathfrak{g}, K)$-cohomology of V [Borel and Wallach 2000, Section I.1.2]. The following two theorems are the essential algebraic ingredients in our proof of (5-13).

Theorem 8.8. *Let V be a Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module with generalized infinitesimal character χ. Let W be a finite-dimensional $\mathfrak{g}_{\mathbb{C}}$-module with infinitesimal character. Let χ^{W^*} be the infinitesimal character of W^*. If $\chi \neq \chi^{W^*}$, then*

$$H^\bullet(\mathfrak{g}, K; V \otimes W) = 0. \tag{8-49}$$

Proof. If χ is the infinitesimal character of V, then (8-49) is a consequence of [Borel and Wallach 2000, Theorem I.5.3(ii)].

In general, let

$$V = V_{n_1} \supset V_{n_1-1} \supset \cdots \supset V_0 \supset V_{-1} = 0 \tag{8-50}$$

be the composition series of V. Then for $0 \leqslant i \leqslant n_1$, we have V_i/V_{i-1} is an irreducible Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-module with infinitesimal character χ. Therefore, for all $0 \leqslant i \leqslant n_1$, we have

$$H^\bullet(\mathfrak{g}, K; (V_i/V_{i-1}) \otimes W) = 0. \tag{8-51}$$

We will show by induction that, for all $0 \leqslant i \leqslant n_1$,

$$H^\bullet(\mathfrak{g}, K; V_i \otimes W) = 0. \tag{8-52}$$

By (8-51), equation (8-52) holds for $i = 0$. Assume that (8-52) holds for some i with $0 \leqslant i \leqslant n_1$. Using the short exact sequence of Harish-Chandra $(\mathfrak{g}_{\mathbb{C}}, K)$-modules

$$0 \to V_i \to V_{i+1} \to V_{i+1}/V_i \to 0, \tag{8-53}$$

we get the long exact sequence of cohomologies

$$\cdots \to H^j(\mathfrak{g}, K; V_i \otimes W) \to H^j(\mathfrak{g}, K; V_{i+1} \otimes W) \to H^j(\mathfrak{g}, K; (V_{i+1}/V_i) \otimes W) \to \cdots . \tag{8-54}$$

By (8-51), (8-54) and by the induction hypotheses, (8-52) holds for $i+1$, which completes the proof of (8-52). □

We denote by $\widehat{G}_u$ the unitary dual of G, that is, the set of equivalence classes of complex irreducible unitary representations π of G on Hilbert spaces V_π. If $(\pi, V_\pi) \in \widehat{G}_u$, by [Knapp 1986, Theorem 8.1], π is irreducible admissible. Let χ_π be the corresponding infinitesimal character.

Theorem 8.9. *If* $(\pi, V_\pi) \in \widehat{G}_u$, *then*

$$\chi_\pi \neq 0 \iff H^\bullet(\mathfrak{g}, K; V_{\pi,K}) = 0. \tag{8-55}$$

Proof. The direction "$\Longrightarrow$" of (8-55) is (8-49). The direction "$\Longleftarrow$" of (8-55) is a consequence of [Vogan and Zuckerman 1984; Vogan 1984; Salamanca-Riba 1999]. Indeed, the irreducible unitary representations with nonvanishing $(\mathfrak{g}, K)$-cohomology are classified and constructed in [Vogan and Zuckerman 1984; Vogan 1984]. By [Salamanca-Riba 1999], the irreducible unitary representations with vanishing infinitesimal character are in the class specified by Vogan and Zuckerman, which implies that their $(\mathfrak{g}, K)$-cohomology are nonvanishing. □

Remark 8.10. The condition that π is unitary is crucial in (8-55). See [Wallach 1988, Section 9.8.3] for a counterexample.

8A5. *The Hecht–Schmid character formula.* Let us recall the main result of [Hecht and Schmid 1983]. Let $Q \subset G$ be a standard parabolic subgroup of G with Lie algebra $\mathfrak{q} \subset \mathfrak{g}$. Let

$$Q = M_Q A_Q N_Q, \quad \mathfrak{q} = \mathfrak{m}_\mathfrak{q} \oplus \mathfrak{a}_\mathfrak{q} \oplus \mathfrak{n}_\mathfrak{q} \tag{8-56}$$

be the corresponding Langlands decompositions [Knapp 1986, Section V.5].

Put $\Delta^+(\mathfrak{a}_\mathfrak{q}, \mathfrak{n}_\mathfrak{q})$ to be the set of all linear forms $\alpha \in \mathfrak{a}_\mathfrak{q}^*$ such that there exists a nonzero element $Y \in \mathfrak{n}_\mathfrak{q}$ such that, for all $a \in \mathfrak{a}_\mathfrak{q}$,

$$\mathrm{ad}(a)Y = \langle \alpha, a\rangle Y. \tag{8-57}$$

Set

$$\mathfrak{a}_\mathfrak{q}^- = \{a \in \mathfrak{a}_\mathfrak{q} : \langle \alpha, a\rangle < 0 \text{ for all } \alpha \in \Delta^+(\mathfrak{a}, \mathfrak{n})\}. \tag{8-58}$$

Put $(M_Q A_Q)^-$ to be the interior in $M_Q A_Q$ of the set

$$\{g \in M_Q A_Q : \det(1 - \mathrm{Ad}(g e^a))|_{\mathfrak{n}_\mathfrak{q}} \geqslant 0 \text{ for all } a \in \mathfrak{a}_\mathfrak{q}^-\}. \tag{8-59}$$

If V is a Harish-Chandra $(\mathfrak{g}_\mathbb{C}, K)$-module, let $H_\bullet(\mathfrak{n}_\mathfrak{q}, V)$ be the $\mathfrak{n}_\mathfrak{q}$-homology of V. By [Hecht and Schmid 1983, Proposition 2.24], $H_\bullet(\mathfrak{n}_\mathfrak{q}, V)$ is a Harish-Chandra $(\mathfrak{m}_{\mathfrak{q}\mathbb{C}} \oplus \mathfrak{a}_{\mathfrak{q}\mathbb{C}}, K \cap M_Q)$-module. We denote by $\Theta^{M_Q A_Q}_{H_\bullet(\mathfrak{n}_\mathfrak{q}, V)}$ the corresponding global character. Also, $M_Q A_Q$ acts on $\mathfrak{n}_\mathfrak{q}$. We denote by $\Theta^{M_Q A_Q}_{\Lambda^\bullet(\mathfrak{n}_\mathfrak{q})}$ the character of $\Lambda^\bullet(\mathfrak{n}_\mathfrak{q})$. By [Hecht and Schmid 1983, Theorem 3.6], the following identity of analytic functions on $(M_Q A_Q)^- \cap G'$ holds:

$$\Theta^G_V|_{(M_Q A_Q)^- \cap G'} = \frac{\sum_{i=0}^{\dim \mathfrak{n}_\mathfrak{q}} (-1)^i \Theta^{M_Q A_Q}_{H_i(\mathfrak{n}_\mathfrak{q}, V)}}{\sum_{i=0}^{\dim \mathfrak{n}_\mathfrak{q}} (-1)^i \Theta^{M_Q A_Q}_{\Lambda^i(\mathfrak{n}_\mathfrak{q})}}\Bigg|_{(M_Q A_Q)^- \cap G'}. \tag{8-60}$$

Take a θ-stable Cartan subalgebra $\mathfrak{h}^{\mathfrak{m}_\mathfrak{q}}$ of $\mathfrak{m}_\mathfrak{q}$. Set $\mathfrak{h}_\mathfrak{q} = \mathfrak{h}^{\mathfrak{m}_\mathfrak{q}} \oplus \mathfrak{a}_\mathfrak{q}$. Then $\mathfrak{h}_\mathfrak{q}$ is a θ-stable Cartan subalgebra of both $\mathfrak{m}_\mathfrak{q} \oplus \mathfrak{a}_\mathfrak{q}$ and $\mathfrak{g}$. Put $\mathfrak{u}_\mathfrak{q}$ to be the compact form of $\mathfrak{m}_\mathfrak{q} \oplus \mathfrak{a}_\mathfrak{q}$. Then $\mathfrak{h}_{\mathfrak{q}\mathbb{R}}$, the real form of $\mathfrak{h}_\mathfrak{q}$, is a Cartan subalgebra of both $\mathfrak{u}_\mathfrak{q}$ and $\mathfrak{u}$. The real root system of $\Delta(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u}_\mathfrak{q})$ is a subset of $\Delta(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u})$ consisting of the elements whose restrictions to $\mathfrak{a}_\mathfrak{q}$ vanish. The set of positive real roots $\Delta^+(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u}) \subset \Delta(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u})$ determines a set of positive real roots $\Delta^+(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u}_\mathfrak{q}) \subset \Delta(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u}_\mathfrak{q})$. Let $\rho_\mathfrak{q}^{\mathfrak{u}}$ and $\rho_\mathfrak{q}^{\mathfrak{u}_\mathfrak{q}}$ be the corresponding half sums of positive real roots.

If V possesses an infinitesimal character with Harish-Chandra parameter $\Lambda \in \mathfrak{h}_{\mathfrak{q}\mathbb{C}}^*$, by [Hecht and Schmid 1983, Corollary 3.32], $H_\bullet(\mathfrak{n}_\mathfrak{q}, V)$ can be decomposed in the sense of (8-4), where the generalized infinitesimal characters are given by

$$\chi_{w\Lambda+\rho_\mathfrak{q}^{\mathfrak{u}}-\rho_\mathfrak{q}^{\mathfrak{u}_\mathfrak{q}}} \tag{8-61}$$

for some $w \in W(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u})$.

Also, $H_\bullet(\mathfrak{n}, V)$ is a Harish-Chandra $(\mathfrak{m}_{\mathfrak{q}\mathbb{C}}, K \cap M_Q)$-module. For $\nu \in \mathfrak{a}_{\mathfrak{q}\mathbb{C}}^*$, let $H_\bullet(\mathfrak{n}, V)_{[\nu]}$ be the largest submodule of $H_\bullet(\mathfrak{n}, V)$ on which $z - \langle 2\sqrt{-1}\pi\nu, z\rangle$ acts nilpotently for all $z \in \mathfrak{a}_{\mathfrak{q}\mathbb{C}}$. Then,

$$H_\bullet(\mathfrak{n}, V) = \bigoplus_\nu H_\bullet(\mathfrak{n}, V)_{[\nu]}, \tag{8-62}$$

where $\nu = (w\Lambda + \rho_\mathfrak{q}^{\mathfrak{u}} - \rho_\mathfrak{q}^{\mathfrak{u}_\mathfrak{q}})|_{\mathfrak{a}_{\mathfrak{q}\mathbb{C}}}$ for some $w \in W(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u})$. Let $\Theta^{M_Q}_{H_\bullet(\mathfrak{n},V)}$ and $\Theta^{M_Q}_{H_\bullet(\mathfrak{n},V)_{[\nu]}}$ be the corresponding global characters. We have the identities of L^1_{loc}-functions: for $m \in M_Q$ and $a \in \mathfrak{a}_\mathfrak{q}$,

$$\Theta^{M_Q A_Q}_{H_\bullet(\mathfrak{n},V)}(me^a) = \sum_\nu e^{2\sqrt{-1}\pi\langle\nu,a\rangle}\Theta^{M_Q}_{H_\bullet(\mathfrak{n},V)_{[\nu]}}(m), \quad \Theta^{M_Q}_{H_\bullet(\mathfrak{n},V)}(m) = \sum_\nu \Theta^{M_Q}_{H_\bullet(\mathfrak{n},V)_{[\nu]}}(m), \tag{8-63}$$

where $\nu = (w\Lambda + \rho_\mathfrak{q}^{\mathfrak{u}} - \rho_\mathfrak{q}^{\mathfrak{u}_\mathfrak{q}})|_{\mathfrak{a}_{\mathfrak{q}\mathbb{C}}}$ for some $w \in W(\mathfrak{h}_{\mathfrak{q}\mathbb{R}}, \mathfrak{u})$.

Suppose now G has compact center and is such that $\delta(G) = 1$. Use the notation in Section 6A. Take $\mathfrak{q} = \mathfrak{m} \oplus \mathfrak{b} \oplus \mathfrak{n}$, and let $Q = M_Q A_Q N_Q \subset G$ be the corresponding parabolic subgroup. Then M is the connected component of the identity in M_Q. Since $K \cap M_Q$ has a finite number of connected components, $H_\bullet(\mathfrak{n}, V)$ is still a Harish-Chandra $(\mathfrak{m}_\mathbb{C} \oplus \mathfrak{b}_\mathbb{C}, K_M)$-module. Also, it is a Harish-Chandra $(\mathfrak{m}_\mathbb{C}, K_M)$-module. Let $\Theta^{MA_Q}_{H_\bullet(\mathfrak{n},V)}$ and $\Theta^{M}_{H_\bullet(\mathfrak{n},V)}$ be the respective global characters.

Recall that $H = \exp(\mathfrak{b})T \subset MA_Q$ is the Cartan subgroup of MA_Q.

Proposition 8.11. *We have*

$$\bigcup_{g\in MA_Q} gH'g^{-1} \subset (M_Q A_Q)^- \cap G'. \tag{8-64}$$

Proof. Put $L' = \bigcup_{g\in MA_Q} gH'g^{-1} \subset MA_Q \cap G'$. Then L' is an open subset of MA_Q. It is enough to show that L' is a subset of (8-59).

By (6-19) and (6-22), for $\gamma = e^a k^{-1} \in H$ with $a \in \mathfrak{b}$ and $k \in T$, we have $\det(1 - \mathrm{Ad}(\gamma))|_\mathfrak{n} \geqslant 0$. Therefore, L' is a subset of (8-59). ☐

8B. ***Formulas for $r_{\eta,\rho}$ and r_j.*** Recall that $\hat{p} : \Gamma\backslash G \to Z$ is the natural projection. The group G acts unitarily on the right on $L^2(\Gamma\backslash G, \hat{p}^*F)$. By [Gel'fand et al. 1969, p. 23, Theorem], we can decompose

$L^2(\Gamma\backslash G, \hat{p}^* F)$ into a direct sum of unitary representations of G,

$$L^2(\Gamma\backslash G, \hat{p}^* F) = \bigoplus_{\pi \in \widehat{G}_u}^{\mathrm{Hil}} n_\rho(\pi) V_\pi, \tag{8-65}$$

with $n_\rho(\pi) < \infty$.

Recall that τ is a real finite-dimensional orthogonal representation of K on the real Euclidean space E_τ, and that $C^{\mathfrak{g},Z,\tau,\rho}$ is the Casimir element of G acting on $C^\infty(Z, \mathcal{F}_\tau \otimes_{\mathbb{C}} F)$. By (8-65), we have

$$\ker C^{\mathfrak{g},Z,\tau,\rho} = \bigoplus_{\substack{\pi \in \widehat{G}_u \\ \chi_\pi(C^{\mathfrak{g}})=0}} n_\rho(\pi)(V_{\pi,K} \otimes_{\mathbb{R}} E_\tau)^K. \tag{8-66}$$

By the properties of elliptic operators, the sum on right-hand side of (8-66) is finite.

We will give two applications of (8-66). In our first application, we take $E_\tau = \Lambda^\bullet(\mathfrak{p}^*)$.

Proposition 8.12. *We have*

$$H^\bullet(Z, F) = \bigoplus_{\substack{\pi \in \widehat{G}_u \\ \chi_\pi=0}} n_\rho(\pi) H^\bullet(\mathfrak{g}, K; V_{\pi,K}). \tag{8-67}$$

If $H^\bullet(Z, F) = 0$, *then for any* $\pi \in \widehat{G}_u$ *such that* $\chi_\pi = 0$, *we have*

$$n_\rho(\pi) = 0. \tag{8-68}$$

Proof. By Hodge theory, and by (4-24) and (8-66), we have

$$H^\bullet(Z, F) = \bigoplus_{\substack{\pi \in \widehat{G}_u \\ \chi_\pi(C^{\mathfrak{g}})=0}} n_\rho(\pi)(V_{\pi,K} \otimes_{\mathbb{R}} \Lambda^\bullet(\mathfrak{p}^*))^K. \tag{8-69}$$

By Hodge theory for Lie algebras [Borel and Wallach 2000, Proposition II.3.1], if $\chi_\pi(C^{\mathfrak{g}}) = 0$, we have

$$(V_{\pi,K} \otimes_{\mathbb{R}} \Lambda^\bullet(\mathfrak{p}^*))^K = H^\bullet(\mathfrak{g}, K; V_{\pi,K}). \tag{8-70}$$

From (8-69) and (8-70), we get

$$H^\bullet(Z, F) = \bigoplus_{\substack{\pi \in \widehat{G}_u \\ \chi_\pi(C^{\mathfrak{g}})=0}} n_\rho(\pi) H^\bullet(\mathfrak{g}, K; V_{\pi,K}). \tag{8-71}$$

By (8-49) and (8-71), we get (8-67).

By Theorem 8.9, and by (8-67), we get (8-68). □

Remark 8.13. Equation (8-67) is [Borel and Wallach 2000, Proposition VII.3.2]. When ρ is a trivial representation, (8-71) is originally due to Matsushima [1967].

In the rest of this section, G is assumed to have compact center and satisfy $\delta(G) = 1$. Recall that η is a real finite-dimensional representation of M satisfying Assumption 7.1, and that $\hat{\eta}$ is defined in (7-1). In our second application of (8-66), we take $\tau = \hat{\eta}$.

Theorem 8.14. *If* $(\pi, V_\pi) \in \widehat{G}_u$, *then*

$$\dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^+)^K - \dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^-)^K = \frac{1}{\chi(K/K_M)} \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} \sum_{j=0}^{2l} (-1)^{i+j} \dim_{\mathbb{C}} H^i\big(\mathfrak{m}, K_M; H_j(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} E_\eta\big). \tag{8-72}$$

Proof. Let $\Lambda(\pi) \in \mathfrak{h}_{\mathbb{C}}^*$ be the Harish-Chandra parameter of the infinitesimal character of π. By (8-29), for $t > 0$, we have

$$\dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^+)^K - \dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^-)^K = \operatorname{vol}(K)^{-1} e^{t\chi_\pi(C^{\mathfrak{g}})/2} \int_{g \in G} \Theta_\pi^G(g) \operatorname{Tr_s}[p_t^{X,\hat{\eta}}(g)]\, dv_G. \tag{8-73}$$

By (7-10), by Proposition 8.7 and by $H \cap K = T$, we have

$$\int_G \Theta_\pi^G(g) \operatorname{Tr_s}[p_t^{X,\hat{\eta}}(g)]\, dv_G = \frac{\operatorname{vol}(T\backslash K)}{|W(H,G)|} \int_{\gamma \in H'} \Theta_\pi^G(\gamma) \operatorname{Tr_s}^{[\gamma]}[\exp(-tC^{\mathfrak{g},X,\hat{\eta}}/2)] \big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}}\big|\, dv_H. \tag{8-74}$$

Since $\gamma = e^a k^{-1} \in H'$ implies $T = K_M(k) = M^0(k)$, by (7-9), (8-73) and (8-74), we have

$$\begin{aligned}
&\dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^+)^K - \dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^-)^K \\
&\quad = \frac{1}{|W(H,G)| \operatorname{vol}(T)} \frac{1}{\sqrt{2\pi t}} \exp\Big(\frac{t}{16} \operatorname{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}] - \frac{t}{2} C^{\mathfrak{u}_{\mathfrak{m}},\eta} + \frac{t}{2}\chi_\pi(C^{\mathfrak{g}})\Big) \\
&\qquad \times \int_{\gamma = e^a k^{-1} \in H'} \Theta_\pi^G(\gamma) \exp(-|a|^2/(2t)) \operatorname{Tr}^{E_\eta}[\eta(k^{-1})] \frac{\big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}}\big|}{\big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\big|^{\frac{1}{2}}}\, dv_H.
\end{aligned} \tag{8-75}$$

By (6-19), for $\gamma = e^a k^{-1} \in H'$, we have

$$\frac{\big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{g}/\mathfrak{h}}\big|}{\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{n}} \big|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{z}_0^\perp}\big|^{\frac{1}{2}}} = e^{-l\langle\alpha,a\rangle} \big|\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{m}/\mathfrak{t}}\big|. \tag{8-76}$$

By (8-60), (8-64), (8-75) and (8-76), we have

$$\begin{aligned}
&\dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^+)^K - \dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^-)^K \\
&\quad = \frac{1}{|W(H,G)| \operatorname{vol}(T)} \frac{1}{\sqrt{2\pi t}} \exp\Big(\frac{t}{16} \operatorname{Tr}^{\mathfrak{u}^\perp(\mathfrak{b})}[C^{\mathfrak{u}(\mathfrak{b}),\mathfrak{u}^\perp(\mathfrak{b})}] - \frac{t}{2} C^{\mathfrak{u}_{\mathfrak{m}},\eta} + \frac{t}{2}\chi_\pi(C^{\mathfrak{g}})\Big) \\
&\qquad \times \sum_{j=0}^{2l} (-1)^j \int_{\gamma = e^a k^{-1} \in H'} \Theta_{H_j(\mathfrak{n}, V_{\pi,K})}^{MA_Q}(\gamma) \exp(-|a|^2/(2t) - l\langle\alpha,a\rangle) \\
&\qquad\qquad \operatorname{Tr}^{E_\eta}[\eta(k^{-1})] \big|\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{m}/\mathfrak{t}}\big|\, dv_H.
\end{aligned} \tag{8-77}$$

By (8-16), there exist $C > 0$ and $c > 0$ such that, for $\gamma = e^a k^{-1} \in H'$, we have

$$|\Theta_{H_j(\mathfrak{n}, V_{\pi,K})}^{MA_Q}(\gamma)| \big|\det(1-\operatorname{Ad}(k^{-1}))|_{\mathfrak{m}/\mathfrak{t}}\big|^{\frac{1}{2}} \leqslant C e^{c|a|}. \tag{8-78}$$

By (8-63), (8-77), (8-78), and by letting $t \to 0$, we get

$$
\begin{aligned}
&\dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^{+})^{K} - \dim_{\mathbb{C}}(V_{\pi,K} \otimes_{\mathbb{R}} \hat{\eta}^{-})^{K} \\
&\quad = \frac{1}{|W(H,G)|\operatorname{vol}(T)} \sum_{j=0}^{2l} (-1)^j \int_{\gamma \in T'} \Theta^{M}_{H_j(\mathfrak{n}, V_{\pi,K})}(\gamma) \operatorname{Tr}^{E_\eta}[\eta(\gamma)] \left|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{m}/\mathfrak{t}}\right| dv_T, \qquad (8\text{-}79)
\end{aligned}
$$

where T' is the set of the regular elements of M in T.

We claim that, for $0 \leqslant j \leqslant 2l$, we have

$$
\begin{aligned}
&\sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} (-1)^i \dim_{\mathbb{C}}\big(H_j(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} \Lambda^i(\mathfrak{p}^*_{\mathfrak{m}}) \otimes_{\mathbb{R}} E_\eta\big)^{K_M} \\
&\qquad = \frac{1}{|W(T,M)|\operatorname{vol}(T)} \int_{\gamma \in T'} \Theta^{M}_{H_j(\mathfrak{n}, V_{\pi,K})}(\gamma) \operatorname{Tr}^{E_\eta}[\eta(\gamma)] \left|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{m}/\mathfrak{t}}\right| dv_T. \qquad (8\text{-}80)
\end{aligned}
$$

Indeed, consider $H_j(\mathfrak{n}, V_{\pi,K})$ as a Harish-Chandra $(\mathfrak{m}_{\mathbb{C}}, K_M)$-module. We can decompose $H_j(\mathfrak{n}, V_{\pi,K})$ in the sense of (8-4), where the generalized infinitesimal characters are given by

$$
\chi_{(w\Lambda(\pi)+\rho^{\mathfrak{u}}-\rho^{\mathfrak{u}(\mathfrak{b})})|_{\mathfrak{t}_{\mathbb{C}}}} \qquad (8\text{-}81)
$$

for some $w \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$. Therefore, it is enough to show (8-80) when $H_j(\mathfrak{n}, V_{\pi,K})$ is replaced by any Harish-Chandra $(\mathfrak{m}_{\mathbb{C}}, K_M)$-module with generalized infinitesimal character $\chi_{(w\Lambda(\pi)+\rho^{\mathfrak{u}}-\rho^{\mathfrak{u}(\mathfrak{b})})|_{\mathfrak{t}_{\mathbb{C}}}}$. Let (π^M, V_{π^M}) be a Hilbert globalization of such a Harish-Chandra $(\mathfrak{m}_{\mathbb{C}}, K_M)$-module. As before, let $C^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta}$ be the Casimir element of M acting on $C^\infty(M, \Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}}) \otimes E_\eta)^{K_M}$, and let $p_t^{X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta}(g)$ be the smooth integral kernel of the heat operator $\exp(-tC^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta}/2)$. Remark that by [Bismut et al. 2017, Proposition 8.4], $C^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta} - C^{\mathfrak{m},E_\eta}$ is the Hodge Laplacian on X_M acting on the differential forms with values in the homogeneous flat vector bundle $M \times_{K_M} E_\eta$. Proceeding as in [Bismut 2011, Theorem 7.8.2], if $\gamma \in M$ is semisimple and nonelliptic, we have

$$
\operatorname{Tr}^{[\gamma]}\big[\exp\big(-t(C^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta} - C^{\mathfrak{m},E_\eta})/2\big)\big] = 0. \qquad (8\text{-}82)
$$

Also, if $\gamma = k^{-1} \in K_M$, then

$$
\operatorname{Tr}^{[\gamma]}\big[\exp\big(-t(C^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta} - C^{\mathfrak{m},E_\eta})/2\big)\big] = \operatorname{Tr}^{E_\eta}[\eta(k^{-1})]e(X_M(k), \nabla^{TX_M(k)}). \qquad (8\text{-}83)
$$

Using (8-82), proceeding as in (8-73) and (8-74), we have

$$
\begin{aligned}
&\sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} (-1)^i \dim_{\mathbb{C}}\big(V_{\pi^M} \otimes_{\mathbb{R}} \Lambda^i(\mathfrak{p}^*_{\mathfrak{m}}) \otimes_{\mathbb{R}} E_\eta\big)^{K_M} \\
&\quad = \operatorname{vol}(K_M)^{-1} \exp(t\chi_{\pi^M}(C^{\mathfrak{m}})/2) \int_{g \in M} \Theta^M_{\pi^M}(g) \operatorname{Tr_s}[p_t^{X_M,\Lambda^\bullet(\mathfrak{p}^*_{\mathfrak{m}})\otimes E_\eta}(g)]\, dv_M \\
&\quad = \frac{\exp(t\chi_{\pi^M}(C^{\mathfrak{m}})/2)}{|W(T,M)|\operatorname{vol}(T)} \int_{\gamma \in T'} \Theta^M_{\pi^M}(\gamma) \operatorname{Tr_s}^{[\gamma]}\big[\exp(-tC^{\mathfrak{m},X_M,\Lambda^\bullet(\mathfrak{p}_{\mathfrak{m}})\otimes E_\eta}/2)\big] \left|\det(1-\operatorname{Ad}(\gamma))|_{\mathfrak{m}/\mathfrak{t}}\right| dv_T. \qquad (8\text{-}84)
\end{aligned}
$$

By (8-83), (8-84), proceeding as in (8-75), and letting $t \to 0$, we get the desired equality (8-80).

The Euler formula asserts

$$\sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} (-1)^i \dim_{\mathbb{C}} \big(H_j(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} \Lambda^i(\mathfrak{p}_{\mathfrak{m}}^*) \otimes_{\mathbb{R}} E_\eta\big)^{K_M} = \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} (-1)^i \dim_{\mathbb{C}} H^i\big(\mathfrak{m}, K_M; H_j(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} E_\eta\big). \tag{8-85}$$

By (3-17), we have

$$W(H, G) = W(T, K), \quad W(T, M) = W(T, K_M). \tag{8-86}$$

By (7-5), (8-79), (8-80) and (8-85)–(8-86), we get (8-72). □

Corollary 8.15. *The following identity holds*:

$$r_{\eta,\rho} = \frac{1}{\chi(K/K_M)} \sum_{\substack{\pi\in\widehat{G}_u \\ \chi_\pi(C^{\mathfrak{g}})=0}} n_\rho(\pi) \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} \sum_{j=0}^{2l} (-1)^{i+j} \dim_{\mathbb{C}} H^i\big(\mathfrak{m}, K_M; H_j(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} E_\eta\big). \tag{8-87}$$

Proof. This is a consequence of (7-58), (8-66) and (8-72). □

Remark 8.16. When $G = \mathrm{SO}^0(p, 1)$ with $p \geqslant 3$ odd, the formula (8-87) is compatible with [Juhl 2001, Theorem 3.11].

We will apply (8-87) to η_j. The following proposition allows us to reduce the first sum in (8-87) to the one over $\pi \in \widehat{G}_u$ with $\chi_\pi = 0$.

Proposition 8.17. *Let* $(\pi, V_\pi) \in \widehat{G}_u$. *Assume* $\chi_\pi(C^{\mathfrak{g}}) = 0$ *and*

$$H^\bullet\big(\mathfrak{m}, K_M; H_\bullet(\mathfrak{n}, V_\pi) \otimes_{\mathbb{R}} \Lambda^j(\mathfrak{n}^*)\big) \neq 0. \tag{8-88}$$

Then the infinitesimal character χ_π *vanishes.*

Proof. Recall that $\Lambda(\pi) \in \mathfrak{h}_{\mathbb{C}}^*$ is a Harish-Chandra parameter of π. We need to show that there is $w \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$ such that

$$w\Lambda(\pi) = \rho^{\mathfrak{u}}. \tag{8-89}$$

Let B^* be the bilinear form on $\mathfrak{g}^*$ induced by B. It extends to $\mathfrak{g}_{\mathbb{C}}^*$ and $\mathfrak{u}^*$ in an obvious way. Since $\chi_\pi(C^{\mathfrak{g},\pi}) = 0$, we have

$$B^*(\Lambda(\pi), \Lambda(\pi)) = B^*(\rho^{\mathfrak{u}}, \rho^{\mathfrak{u}}). \tag{8-90}$$

We identify $\mathfrak{h}_{\mathbb{R}}^* = \sqrt{-1}\mathfrak{b}^* \oplus \mathfrak{t}^*$. By definition,

$$\rho^{\mathfrak{u}} = \left(\frac{l\alpha}{2\sqrt{-1}\pi}, \rho^{\mathfrak{u}_{\mathfrak{m}}}\right) \in \sqrt{-1}\mathfrak{b}^* \oplus \mathfrak{t}^* \quad \text{and} \quad \rho^{\mathfrak{u}(\mathfrak{b})} = (0, \rho^{\mathfrak{u}_{\mathfrak{m}}}) \in \sqrt{-1}\mathfrak{b}^* \oplus \mathfrak{t}^*. \tag{8-91}$$

By (8-49), (8-81) and (8-88), there exist $w \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$, $w' \in W(\mathfrak{t}, \mathfrak{u}_{\mathfrak{m}}) \subset W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$ and the highest real weight $\mu_j \in \mathfrak{t}^*$ of an irreducible subrepresentation of $\mathfrak{m}_{\mathbb{C}}$ on $\Lambda^j(\mathfrak{n}_{\mathbb{C}}) \simeq \Lambda^j(\bar{\mathfrak{n}}_{\mathbb{C}}^*)$ such that

$$w\Lambda(\pi)|_{\mathfrak{t}_{\mathbb{C}}} = w'(\mu_j + \rho^{\mathfrak{u}_{\mathfrak{m}}}). \tag{8-92}$$

By (6-58), (8-90) and (8-92), there exists $w'' \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$ such that

$$w''\Lambda(\pi) = \left(\pm\frac{(l-j)\alpha}{2\sqrt{-1}\pi}, \mu_j + \rho^{\mathfrak{u}_{\mathfrak{m}}}\right) = \left(\pm\frac{(l-j)\alpha}{2\sqrt{-1}\pi}, \mu_j\right) + \rho^{\mathfrak{u}(\mathfrak{b})}. \tag{8-93}$$

In particular, $w''\Lambda(\pi) \in \mathfrak{h}_{\mathbb{R}}^*$.

Clearly, $\big((j-l)\alpha/(2\sqrt{-1}\pi), \mu_j\big) \in \mathfrak{h}_{\mathbb{R}}^*$ is the highest real weight of an irreducible subrepresentation of $\mathfrak{m}_{\mathbb{C}} \oplus \mathfrak{b}_{\mathbb{C}}$ on $\Lambda^j(\bar{\mathfrak{n}}_{\mathbb{C}}^*) \otimes_{\mathbb{C}} (\det(\mathfrak{n}_{\mathbb{C}}))^{-\frac{1}{2}}$. By (6-33), $\big((j-l)\alpha/(2\sqrt{-1}\pi), \mu_j\big) \in \mathfrak{h}_{\mathbb{R}}^*$ is the highest real weight of an irreducible subrepresentation of $\mathfrak{m}_{\mathbb{C}} \oplus \mathfrak{b}_{\mathbb{C}}$ on $S^{\mathfrak{u}^{\perp}(\mathfrak{b})}$. By [Borel and Wallach 2000, Lemma II.6.9], there exists $w_1 \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$ such that

$$\left(\frac{(j-l)\alpha}{2\sqrt{-1}\pi}, \mu_j\right) = w_1\rho^{\mathfrak{u}} - \rho^{\mathfrak{u}(\mathfrak{b})}. \tag{8-94}$$

Similarly, $\big((l-j)\alpha/(2\sqrt{-1}\pi), \mu_j\big) \in \mathfrak{h}_{\mathbb{R}}^*$ is the highest real weight of an irreducible subrepresentation of $\mathfrak{m}_{\mathbb{C}} \oplus \mathfrak{b}_{\mathbb{C}}$ on both $\Lambda^{2l-j}(\bar{\mathfrak{n}}_{\mathbb{C}}^*) \otimes_{\mathbb{C}} (\det(\mathfrak{n}_{\mathbb{C}}))^{-\frac{1}{2}}$ and $S^{\mathfrak{u}^{\perp}(\mathfrak{b})}$. Therefore, there exists $w_2 \in W(\mathfrak{h}_{\mathbb{R}}, \mathfrak{u})$ such that

$$\left(\frac{(l-j)\alpha}{2\sqrt{-1}\pi}, \mu_j\right) = w_2\rho^{\mathfrak{u}} - \rho^{\mathfrak{u}(\mathfrak{b})}. \tag{8-95}$$

By (8-93)–(8-95), we get (8-89). □

Corollary 8.18. *For $0 \leqslant j \leqslant 2l$, we have*

$$r_j = \frac{1}{\chi(K/K_M)} \sum_{\substack{\pi \in \widehat{G}_u \\ \chi_\pi = 0}} n_\rho(\pi) \sum_{i=0}^{\dim \mathfrak{p}_{\mathfrak{m}}} \sum_{k=0}^{2l} (-1)^{i+k} \dim_{\mathbb{C}} H^i\big(\mathfrak{m}, K_M; H_k(\mathfrak{n}, V_{\pi,K}) \otimes_{\mathbb{R}} \Lambda^j(\mathfrak{n}^*)\big). \tag{8-96}$$

If $H^\bullet(Z, F) = 0$, then for all $0 \leqslant j \leqslant 2l$,

$$r_j = 0. \tag{8-97}$$

Proof. This is a consequence of Proposition 8.12, Corollary 8.15 and Proposition 8.17. □

Remark 8.19. By (7-75) and (8-97), we get (5-13) when G has compact center and $\delta(G) = 1$.

Acknowledgements

I am deeply grateful to Prof. J.-M. Bismut for sharing his insights on this problem. Without his help, this paper would never have been written. I am indebted to Prof. W. Müller for his detailed explanation of the paper [Moscovici and Stanton 1991]. I would also like to thank Prof. L. Clozel and Prof. M. Vergne for very useful discussions. I appreciate the encouragement and constant support of Prof. X. Ma. I am also indebted to referees who read the manuscript very carefully and offered detailed suggestions for its improvement.

This article was prepared while I was visiting Max-Planck-Institut für Mathematik at Bonn and the Humboldt-Universität zu Berlin. I would like to thank both institutions for their hospitality.

This research has been financially supported by Collaborative Research Centre “Space-Time-Matter” (SFB 647), funded by the German Research Foundation (DFG).

Index

References

[Barbasch and Moscovici 1983] D. Barbasch and H. Moscovici, "L^2-index and the Selberg trace formula", *J. Funct. Anal.* **53**:2 (1983), 151–201. MR Zbl

[Berline et al. 2004] N. Berline, E. Getzler, and M. Vergne, *Heat kernels and Dirac operators*, Springer, 2004. MR Zbl

[Bismut 2005] J.-M. Bismut, "The hypoelliptic Laplacian on the cotangent bundle", *J. Amer. Math. Soc.* **18**:2 (2005), 379–476. MR Zbl

[Bismut 2011] J.-M. Bismut, *Hypoelliptic Laplacian and orbital integrals*, Annals of Mathematics Studies **177**, Princeton University Press, 2011. MR Zbl

[Bismut and Goette 2004] J.-M. Bismut and S. Goette, "Equivariant de Rham torsions", *Ann. of Math.* (2) **159**:1 (2004), 53–216. MR Zbl

[Bismut and Zhang 1992] J.-M. Bismut and W. Zhang, *An extension of a theorem by Cheeger and Müller*, Astérisque **205**, Société Mathématique de France, Paris, 1992. MR Zbl

[Bismut et al. 2017] J.-M. Bismut, X. Ma, and W. Zhang, "Asymptotic torsion and Toeplitz operators", *J. Inst. Math. Jussieu* **16**:2 (2017), 223–349. MR Zbl

[Borel and Wallach 2000] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, 2nd ed., Mathematical Surveys and Monographs **67**, Amer. Math. Society, Providence, RI, 2000. MR Zbl

[Bott 1965] R. Bott, "The index theorem for homogeneous differential operators", pp. 167–186 in *Differential and combinatorial topology: a symposium in honor of Marston Morse*, edited by S. S. Cairns, Princeton University Press, 1965. MR Zbl

[Bröcker and tom Dieck 1985] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics **98**, Springer, 1985. MR Zbl

[Cheeger 1979] J. Cheeger, "Analytic torsion and the heat equation", *Ann. of Math.* (2) **109**:2 (1979), 259–322. MR Zbl

[Duistermaat et al. 1979] J. J. Duistermaat, J. A. C. Kolk, and V. S. Varadarajan, "Spectra of compact locally symmetric manifolds of negative curvature", *Invent. Math.* **52**:1 (1979), 27–93. MR Zbl

[Dyatlov and Zworski 2016] S. Dyatlov and M. Zworski, "Dynamical zeta functions for Anosov flows via microlocal analysis", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:3 (2016), 543–577. MR Zbl

[Dyatlov and Zworski 2017] S. Dyatlov and M. Zworski, "Ruelle zeta function at zero for surfaces", *Invent. Math.* (online publication April 2017).

[Fried 1986] D. Fried, "Analytic torsion and closed geodesics on hyperbolic manifolds", *Invent. Math.* **84**:3 (1986), 523–540. MR Zbl

[Fried 1987] D. Fried, "Lefschetz formulas for flows", pp. 19–69 in *The Lefschetz centennial conference, Part III* (Mexico City, 1984), Contemp. Math. **58**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Gel'fand et al. 1969] I. M. Gel'fand, M. I. Graev, and I. I. Pyatetskii-Shapiro, *Representation theory and automorphic functions*, W. B. Saunders Co., Philadelphia, PA, 1969. MR

[Giulietti et al. 2013] P. Giulietti, C. Liverani, and M. Pollicott, "Anosov flows and dynamical zeta functions", *Ann. of Math.* (2) **178**:2 (2013), 687–773. MR Zbl

[Hecht and Schmid 1983] H. Hecht and W. Schmid, "Characters, asymptotics and $\mathfrak{n}$-homology of Harish-Chandra modules", *Acta Math.* **151**:1-2 (1983), 49–151. MR Zbl

[Helgason 1978] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics **80**, Academic Press, New York, 1978. MR Zbl

[Hirzebruch 1966] F. Hirzebruch, *Topological methods in algebraic geometry*, Die Grundlehren der Mathematischen Wissenschaften **131**, Springer, 1966. MR Zbl

[Hitchin 1974] N. Hitchin, "Harmonic spinors", *Advances in Math.* **14** (1974), 1–55. MR Zbl

[Juhl 2001] A. Juhl, *Cohomological theory of dynamical zeta functions*, Progress in Mathematics **194**, Birkhäuser, Basel, 2001. MR Zbl

[Knapp 1986] A. W. Knapp, *Representation theory of semisimple groups: an overview based on examples*, Princeton Mathematical Series **36**, Princeton University Press, 1986. MR Zbl

[Knapp 2002] A. W. Knapp, *Lie groups beyond an introduction*, 2nd ed., Progress in Mathematics **140**, Birkhäuser, Boston, 2002. MR Zbl

[Kobayashi and Nomizu 1963] S. Kobayashi and K. Nomizu, *Foundations of differential geometry, Vol. I*, Interscience Publishers, New York, 1963. MR Zbl

[Kostant 1976] B. Kostant, "On Macdonald's η-function formula, the Laplacian and generalized exponents", *Advances in Math.* **20**:2 (1976), 179–212. MR Zbl

[Li and Yau 1986] P. Li and S.-T. Yau, "On the parabolic kernel of the Schrödinger operator", *Acta Math.* **156**:3-4 (1986), 153–201. MR Zbl

[Ma 2017] X. Ma, "Geometric hypoelliptic Laplacian and orbital integrals (after Bismut, Lebeau and Shen)", lecture given at Séminaire Bourbaki, Exp. No. 1130, March 11 2017, available at http://www.bourbaki.ens.fr/TEXTES/1130.pdf. Video available at https://www.youtube.com/watch?v=dCDwN-HqcJw.

[Ma and Marinescu 2015] X. Ma and G. Marinescu, "Exponential estimate for the asymptotics of Bergman kernels", *Math. Ann.* **362**:3-4 (2015), 1327–1347. MR Zbl

[Matsushima 1967] Y. Matsushima, "A formula for the Betti numbers of compact locally symmetric Riemannian manifolds", *J. Differential Geometry* **1** (1967), 99–109. MR Zbl

[Milnor 1968a] J. Milnor, "A note on curvature and fundamental group", *J. Differential Geometry* **2** (1968), 1–7. MR Zbl

[Milnor 1968b] J. W. Milnor, "Infinite cyclic coverings", pp. 115–133 in *Conference on the Topology of Manifolds* (Lansing, MI, 1967), Prindle, Weber & Schmidt, Boston, 1968. MR Zbl

[Moscovici and Stanton 1991] H. Moscovici and R. J. Stanton, "R-torsion and zeta functions for locally symmetric manifolds", *Invent. Math.* **105**:1 (1991), 185–216. MR Zbl

[Müller 1978] W. Müller, "Analytic torsion and R-torsion of Riemannian manifolds", *Adv. in Math.* **28**:3 (1978), 233–305. MR Zbl

[Quillen 1985] D. Quillen, "Superconnections and the Chern character", *Topology* **24**:1 (1985), 89–95. MR Zbl

[Ray and Singer 1971] D. B. Ray and I. M. Singer, "R-torsion and the Laplacian on Riemannian manifolds", *Advances in Math.* **7** (1971), 145–210. MR Zbl

[Reidemeister 1935] K. Reidemeister, "Homotopieringe und Linsenräume", *Abh. Math. Sem. Univ. Hamburg* **11**:1 (1935), 102–109. MR Zbl

[Salamanca-Riba 1999] S. A. Salamanca-Riba, "On the unitary dual of real reductive Lie groups and the $A_{\mathfrak{g}}(\lambda)$ modules: the strongly regular case", *Duke Math. J.* **96**:3 (1999), 521–546. MR Zbl

[Satake 1957] I. Satake, "The Gauss–Bonnet theorem for V-manifolds", *J. Math. Soc. Japan* **9** (1957), 464–492. MR Zbl

[Seeley 1967] R. T. Seeley, "Complex powers of an elliptic operator", pp. 288–307 in *Singular integrals* (Chicago, IL, 1966), Amer. Math. Soc., Providence, RI, 1967. MR Zbl

[Selberg 1960] A. Selberg, "On discontinuous groups in higher-dimensional symmetric spaces", pp. 147–164 in *Contributions to function theory* (Bombay, 1960), Tata Institute of Fundamental Research, Bombay, 1960. MR Zbl

[Shen 2016] S. Shen, "Analytic torsion, dynamical zeta functions and orbital integrals", *C. R. Math. Acad. Sci. Paris* **354**:4 (2016), 433–436. MR

[Varadarajan 1977] V. S. Varadarajan, *Harmonic analysis on real reductive groups*, Lecture Notes in Mathematics **576**, Springer, 1977. MR Zbl

[Vogan 1984] D. A. Vogan, Jr., "Unitarizability of certain series of representations", *Ann. of Math.* (2) **120**:1 (1984), 141–187. MR Zbl

[Vogan 2008] D. A. Vogan, Jr., "Unitary representations and complex analysis", pp. 259–344 in *Representation theory and complex analysis*, Lecture Notes in Math. **1931**, Springer, 2008. MR Zbl

[Vogan and Zuckerman 1984] D. A. Vogan, Jr. and G. J. Zuckerman, "Unitary representations with nonzero cohomology", *Compositio Math.* **53**:1 (1984), 51–90. MR Zbl

[Voros 1987] A. Voros, "Spectral functions, special functions and the Selberg zeta function", *Comm. Math. Phys.* **110**:3 (1987), 439–465. MR Zbl

[Wallach 1988] N. R. Wallach, *Real reductive groups, I*, Pure and Applied Mathematics **132**, Academic Press, Boston, 1988. MR Zbl

Received 14 Feb 2016. Revised 16 Jun 2017. Accepted 17 Jul 2017.

SHU SHEN: shu.shen@math.u-psud.fr
Département de Mathématique Bâtiment 425, Université Paris-Sud, Orsay, France

mathematical sciences publishers

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.75

EXISTENCE THEOREMS OF THE FRACTIONAL YAMABE PROBLEM

SEUNGHYEOK KIM, MONICA MUSSO AND JUNCHENG WEI

Let X be an asymptotically hyperbolic manifold and M its conformal infinity. This paper is devoted to deducing several existence results of the fractional Yamabe problem on M under various geometric assumptions on X and M. Firstly, we handle when the boundary M has a point at which the mean curvature is negative. Secondly, we re-encounter the case when M has zero mean curvature and satisfies one of the following conditions: nonumbilic, umbilic and a component of the covariant derivative of the Ricci tensor on $\overline{X}$ is negative, or umbilic and nonlocally conformally flat. As a result, we replace the geometric restrictions given by González and Qing (2013) and González and Wang (2017) with simpler ones. Also, inspired by Marques (2007) and Almaraz (2010), we study lower-dimensional manifolds. Finally, the situation when X is Poincaré–Einstein and M is either locally conformally flat or 2-dimensional is covered under a certain condition on a Green's function of the fractional conformal Laplacian.

1. Introduction and the main results

Given $n \in \mathbb{N}$, let X^{n+1} be an $(n+1)$-dimensional smooth manifold with smooth boundary M^n. A function ρ in X is called a *defining function* of the boundary M in X if $\rho > 0$ in X and $\rho = 0$, $d\rho \neq 0$ on M. A metric g^+ in X is *conformally compact* if there exists a boundary-defining function ρ such that the conformal metric $\bar{g} := \rho^2 g^+$ extends to M and the closure $(\overline{X}, \bar{g})$ of X is compact. This induces the conformal class $[\hat{h}]$ of the metric $\hat{h} := \bar{g}|_M$, which is referred to as the *conformal infinity* of (X, g^+). A manifold (X, g^+) is called *asymptotically hyperbolic* if g^+ is conformally compact and $|d\rho|_{\bar{g}} \to 1$ as $\rho \to 0$. Also if (X, g^+) is conformally compact and Einstein, then it is said to be *Poincaré–Einstein* or *conformally compact Einstein*. All Poincaré–Einstein manifolds can be shown to be asymptotically hyperbolic.

Suppose an asymptotically hyperbolic manifold (X, g^+) with the conformal infinity $(M^n, [\hat{h}])$ is given. Also, for any $\gamma \in (0,1)$, let $P^\gamma_{\hat{h}} = P^\gamma[g^+, \hat{h}]$ be the *fractional conformal Laplacian* whose principle symbol is equal to that of $(-\Delta_{\hat{h}})^\gamma$; see [Mazzeo and Melrose 1987; Joshi and Sá Barreto 2000; Graham and Zworski 2003; Chang and González 2011; González and Qing 2013] for its precise definition. In this article, we are interested in finding a conformal metric $\hat{h}$ on M with constant *fractional scalar curvature* $Q^\gamma_{\hat{h}} := P^\gamma_{\hat{h}}(1)$. This problem is called the *fractional Yamabe problem* or the *γ-Yamabe problem*, and it was introduced and investigated by González and Qing [2013] and González and Wang [2017]. By imposing some restrictions on the dimension and geometric behavior of the manifold, the authors obtained existence results when M is nonumbilic or M is umbilic but not locally conformally flat. Here we relieve the hypotheses made in [González and Qing 2013; González and Wang 2017] and examine when the bubble (see (1-13) below for its precise definition) cannot be used as an appropriate test function.

MSC2010: primary 53C21; secondary 35R11, 53A30.
Keywords: fractional Yamabe problem, conformal geometry, existence.

As its name alludes, the fractional conformal Laplacian $P^{\gamma}_{\hat{h}}$ has the *conformal covariance property*: it holds that

$$P^{\gamma}_{\hat{h}_w}(u) = w^{-\frac{n+2\gamma}{n-2\gamma}} P^{\gamma}_{\hat{h}}(wu) \tag{1-1}$$

for a conformal change of the metric $\hat{h}_w = w^{4/(n-2\gamma)}\hat{h}$. Hence the fractional Yamabe problem can be formulated as looking for a positive solution of the nonlocal equation

$$P^{\gamma}_{\hat{h}} u = c u^{\frac{n+2\gamma}{n-2\gamma}} \quad \text{on } M \tag{1-2}$$

for some $c \in \mathbb{R}$ provided $n > 2\gamma$. On the other hand, if (X, g^+) is Poincaré–Einstein, then $P^{\gamma}_{\hat{h}}$ and $Q^{\gamma}_{\hat{h}}$ with $\gamma = 1$ precisely match with the classical *conformal Laplacian* $L_{\hat{h}}$ and a constant multiple of the scalar curvature $R[\hat{h}]$ on $(M, \hat{h})$:

$$P^1_{\hat{h}} = L_{\hat{h}} := -\Delta_{\hat{h}} + \frac{n-2}{4(n-1)} R[\hat{h}] \quad \text{and} \quad Q^1_{\hat{h}} = \frac{n-2}{4(n-1)} R[\hat{h}]. \tag{1-3}$$

If $\gamma = 2$, they coincide with the Paneitz operator [2008] and Branson's Q-curvature [1985] (see [Graham and Zworski 2003, Proposition 4.3] for its proof). Hence, in this case, the 1- and 2-Yamabe problems are reduced to the classical Yamabe problem and the Q-curvature problem, respectively.

Thanks to the efforts of various mathematicians, a complete solution of the Yamabe problem is known. After Yamabe [1960] raised the problem and suggested an outline of the proof, Trudinger [1968] first obtained a least energy solution to (1-2) under the setting that the scalar curvature of $(M, \hat{h})$ is nonpositive. Successively, Aubin [1976] examined the case when $n \geq 6$ and M is nonlocally conformally flat, and Schoen [1984] gave an affirmative answer when $n = 3, 4, 5$ or M is locally conformally flat by using the positive mass theorem [Schoen and Yau 1979a; 1979b; 1988]. Lee and Parker [1987] provided a new proof which unified the local proof of Aubin and the global proof of Schoen, introducing the notion of the conformal normal coordinates.

There also have been lots of results on the Q-curvature problem ($\gamma = 2$) for 4-dimensional manifolds $(M^4, [\hat{h}])$. By the Chern–Gauss–Bonnet formula, the *total Q-curvature*

$$k_P := \int_{M^4} Q^2_{\hat{h}} \, dv_{\hat{h}},$$

where $dv_{\hat{h}}$ is the volume form of $(M, \hat{h})$, is a conformal invariant. Gursky [1999] proved that if a manifold M^4 has the positive Yamabe constant $\Lambda^1(M, [\hat{h}]) > 0$, see (1-10), and satisfies $k_P \geq 0$, then its Paneitz operator $P^2_{\hat{h}}$ has the properties

$$\ker P^2_{\hat{h}} = \mathbb{R} \quad \text{and} \quad P^2_{\hat{h}} \geq 0. \tag{1-4}$$

Also Chang and Yang [1995] proved that any compact 4-manifold such that (1-4) and $k_P < 8\pi^2$ hold has a solution to

$$P^2_{\hat{h}} u + 2Q^2_{\hat{h}} u = 2ce^{4u} \quad \text{on } M,\ c \in \mathbb{R},$$

where $Q^2_{\hat{h}}$ is the Q-curvature. This result was generalized by Djadli and Malchiodi [2008] where only $\ker P^2_{\hat{h}} = \mathbb{R}$ and $k_P \neq 8m\pi^2$ for all $m \in \mathbb{N}$ are demanded. For other dimensions than 4, Gursky and

Malchiodi [2015] recently discovered the strong maximum principle of $P^2_{\hat{h}}$ for manifolds M^n ($n \geq 5$) with nonnegative scalar curvature and semipositive Q-curvature. Motivated by this result, Hang and Yang developed the existence theory of (1-2) for a general class of manifolds M^n, including ones such that $\Lambda^1(M, [\hat{h}]) > 0$ and there exists $\hat{h}' \in [\hat{h}]$ with $Q^2_{\hat{h}'} > 0$, provided $n \geq 5$ [Hang and Yang 2015; 2016b] or $n = 3$ [Hang and Yang 2004; 2015; 2016a]. In [Hang and Yang 2016b], the positive mass theorem for the Paneitz operator [Humbert and Raulot 2009; Gursky and Malchiodi 2015] was used to construct a test function. We also point out that a solution to (1-2) was obtained in [Qing and Raske 2006] for a locally conformally flat manifold ($n \geq 5$) with positive Yamabe constant and Poincaré exponent less than $(n-4)/2$.

In addition, when $\gamma = \frac{1}{2}$, the fractional Yamabe problem has a deep relationship with the boundary Yamabe problem proposed by Cherrier [1984] and Escobar [1992a], which can be regarded as a generalization of the Riemann mapping theorem: It asks if a compact manifold $\overline{X}$ with boundary is conformally equivalent to one of zero scalar curvature whose boundary M has constant mean curvature. It was solved by the series of works by Escobar [1992a; 1996], Marques [2005; 2007] and Almaraz [2010] who used a minimization argument. See also [Chen 2009; Mayer and Ndiaye 2015a], in which different approaches are pursued. It is worthwhile to mention that there is another type of boundary Yamabe problem also suggested by Escobar [1992b]: find a conformal metric such that the scalar curvature of X is constant and the boundary M is minimal. It was further studied by Brendle and Chen [2014] and Mayer and Ndiaye [2015b].

Chang and González [2011] observed that the fractional conformal Laplacian, defined through scattering theory (see, e.g., [Mazzeo and Melrose 1987; Joshi and Sá Barreto 2000; Graham and Zworski 2003]), can be described in terms of Dirichlet–Neumann operators; see also [Case and Chang 2016]. Specifically, (1-2) has an equivalent extension problem, which is degenerate elliptic but local.

Theorem A. *Suppose that $n > 2\gamma$, $\gamma \in (0,1)$, and (X, g^+) is an asymptotically hyperbolic manifold with conformal infinity $(M, [\hat{h}])$. Assume also that ρ is a defining function associated to M such that $|d\rho|_{\bar{g}} = 1$ near M (such ρ is called geodesic), and $\bar{g} = \rho^2 g^+$ is a metric of the compact manifold $\overline{X}$. In addition, we let the mean curvature H on $(M, \hat{h}) \subset (\overline{X}, \bar{g})$ be 0 if $\gamma \in \left(\frac{1}{2}, 1\right)$, and set*

$$E(\rho) = \rho^{-1-s}(-\Delta_{g^+} - s(n-s))\rho^{n-s} \quad \text{in } X, \tag{1-5}$$

where $s := n/2 + \gamma$. It can be shown that (1-5) is reduced to

$$E(\rho) = \frac{n-2\gamma}{4n}\left[R[\bar{g}] - (n(n+1) + R[g^+])\rho^{-2}\right]\rho^{1-2\gamma} \quad \text{near } M, \tag{1-6}$$

where $R[\bar{g}]$ and $R[g^+]$ are the scalar curvature of $(\overline{X}, \bar{g})$ and (X, g^+), respectively.

(1) *If a positive function U satisfies*

$$\begin{cases} -\operatorname{div}_{\bar{g}}(\rho^{1-2\gamma}\nabla U) + E(\rho)U = 0 & \text{in } (X, \bar{g}), \\ U = u & \text{on } M \end{cases} \tag{1-7}$$

and

$$\partial^\gamma_\nu U := -\kappa_\gamma\left(\lim_{\rho\to 0+} \rho^{1-2\gamma}\frac{\partial U}{\partial \rho}\right) = \begin{cases} c u^{\frac{n+2\gamma}{n-2\gamma}} & \text{for } \gamma \in (0,1) \setminus \left\{\frac{1}{2}\right\}, \\ c u^{\frac{n+2\gamma}{n-2\gamma}} - \left(\frac{1}{2}(n-1)\right)Hu & \text{for } \gamma = \left\{\frac{1}{2}\right\} \end{cases} \tag{1-8}$$

on M, then u solves (1-2). *Here $\kappa_\gamma > 0$ is the constant whose explicit value is given in* (1-23) *below and ν stands for the outward unit normal vector with respect to the boundary M.*

(2) *Assume further that the first L^2-eigenvalue $\lambda_1(-\Delta_{g^+})$ of the Laplace-Beltrami operator $-\Delta_{g^+}$ satisfies*

$$\lambda_1(-\Delta_{g^+}) > \tfrac{1}{4}n^2 - \gamma^2. \tag{1-9}$$

Then there is a special defining function ρ^* such that $E(\rho^*) = 0$ in X and $\rho^*(\rho) = \rho\,(1 + O(\rho^{2\gamma}))$ near M. Furthermore the function $\widetilde{U} := (\rho/\rho^*)^{(n-2\gamma)/2}U$ solves a degenerate elliptic equation of pure divergent form

$$\begin{cases} -\operatorname{div}_{\bar{g}^*}((\rho^*)^{1-2\gamma}\nabla\widetilde{U}) = 0 & \text{in } (X, \bar{g}^*), \\ \partial_\nu^\gamma \widetilde{U} = -\kappa_\gamma \left(\lim\limits_{\rho^*\to 0+} (\rho^*)^{1-2\gamma} \dfrac{\partial \widetilde{U}}{\partial \rho^*} \right) = P_{\hat{h}}^\gamma u - Q_{\hat{h}}^\gamma u = c u^{\frac{n+2\gamma}{n-2\gamma}} - Q_{\hat{h}}^\gamma u & \text{on } M, \end{cases}$$

where $\bar{g}^* := (\rho^*)^2 g^+$ and $Q_{\hat{h}}^\gamma$ is the fractional scalar curvature.

Notice that in order to seek a solution of (1-2), it is natural to introduce the *γ-Yamabe functional*

$$I_{\hat{h}}^\gamma[u] = \frac{\int_M u P_{\hat{h}}^\gamma u \, dv_{\hat{h}}}{\left(\int_M |u|^{\frac{2n}{n-2\gamma}} \, dv_{\hat{h}}\right)^{\frac{n-2\gamma}{n}}} \quad \text{for } u \in H^\gamma(M) \setminus \{0\}, \tag{1-10}$$

where $H^\gamma(M)$ denotes the standard fractional Sobolev space, and its infimum $\Lambda^\gamma(M, [\hat{h}])$, called the *$\gamma$-Yamabe constant*. By the previous theorem and the energy inequality due to Case [2017, Theorem 1.1], it follows under the assumption (1-9) that if one defines the functionals

$$\bar{I}_{\hat{h}}^\gamma[U] = \frac{\kappa_\gamma \int_X (\rho^{1-2\gamma}|\nabla U|_{\bar{g}}^2 + E(\rho)U^2) \, dv_{\bar{g}}}{\left(\int_M |U|^{\frac{2n}{n-2\gamma}} \, dv_{\hat{h}}\right)^{\frac{n-2\gamma}{n}}}, \quad \tilde{I}_{\hat{h}}^\gamma[U] = \frac{\kappa_\gamma \int_X (\rho^*)^{1-2\gamma}|\nabla U|_{\bar{g}}^2 \, dv_{\bar{g}} + \int_M Q_{\hat{h}}^\gamma U^2 \, dv_{\hat{h}}}{\left(\int_M |U|^{\frac{2n}{n-2\gamma}} \, dv_{\hat{h}}\right)^{\frac{n-2\gamma}{n}}} \tag{1-11}$$

for each element U of the weighted Sobolev space $W^{1,2}(X, \rho^{1-2\gamma})$ such that $U \neq 0$ on M (in view of (1-8), a suitable modification is necessary if $\gamma = \frac{1}{2}$), and values

$$\begin{aligned} \bar{\Lambda}^\gamma(X, [\hat{h}]) &= \inf\left\{\bar{I}_{\hat{h}}^\gamma[U] : U \in W^{1,2}(X, \rho^{1-2\gamma}), U \neq 0 \text{ on } M\right\}, \\ \tilde{\Lambda}^\gamma(X, [\hat{h}]) &= \inf\left\{\tilde{I}_{\hat{h}}^\gamma[U] : U \in W^{1,2}(X, \rho^{1-2\gamma}), U \neq 0 \text{ on } M\right\}, \end{aligned}$$

then

$$\Lambda^\gamma(M, [\hat{h}]) = \bar{\Lambda}^\gamma(X, [\hat{h}]) = \tilde{\Lambda}^\gamma(X, [\hat{h}]) > -\infty.$$

Besides it was shown in [González and Qing 2013] that the sign of c in (1-2) is the same as that of $\Lambda^\gamma(M, [\hat{h}])$, as in the local case $\gamma = 1$.

On the other hand, the Sobolev trace inequality

$$\left(\int_{\mathbb{R}^n} |U(\bar{x}, 0)|^{\frac{2n}{n-2\gamma}} \, d\bar{x}\right)^{\frac{n-2\gamma}{n}} \leq S_{n,\gamma} \int_0^\infty \int_{\mathbb{R}^n} x_{n+1}^{1-2\gamma} |\nabla U(\bar{x}, x_n)|^2 \, d\bar{x} \, dx_{n+1} \tag{1-12}$$

is true for all functions U which belong to the homogeneous weighted Sobolev space $D^{1,2}(\mathbb{R}^{n+1}_+, x_{n+1}^{1-2\gamma})$. In addition, the equality is attained by $U = cW_{\lambda,\sigma}$ for any $c \in \mathbb{R}$, $\lambda > 0$ and $\sigma \in \mathbb{R}^n = \partial\mathbb{R}^{n+1}_+$, where $W_{\lambda,\sigma}$ are the *bubbles* defined as

$$\begin{aligned} W_{\lambda,\sigma}(\bar{x}, x_{n+1}) &= p_{n,\gamma} \int_{\mathbb{R}^n} \frac{x_{n+1}^{2\gamma}}{(|\bar{x}-\bar{y}|^2 + x_{n+1}^2)^{\frac{n+2\gamma}{2}}} w_{\lambda,\sigma}(\bar{y})\, d\bar{y} \\ &= g_{n,\gamma} \int_{\mathbb{R}^n} \frac{1}{(|\bar{x}-\bar{y}|^2 + x_{n+1}^2)^{\frac{n-2\gamma}{2}}} w_{\lambda,\sigma}^{\frac{n+2\gamma}{n-2\gamma}}(\bar{y})\, d\bar{y} \end{aligned} \tag{1-13}$$

with

$$w_{\lambda,\sigma}(\bar{x}) := \alpha_{n,\gamma} \left(\frac{\lambda}{\lambda^2 + |\bar{x}-\sigma|^2} \right)^{\frac{n-2\gamma}{2}} = W_{\lambda,\sigma}(\bar{x}, 0). \tag{1-14}$$

The values of the positive numbers $p_{n,\gamma}$, $g_{n,\gamma}$ and $\alpha_{n,\gamma}$ can be found in (1-23). Particularly, it holds that

$$\begin{cases} -\operatorname{div}(x_{n+1}^{1-2\gamma} \nabla W_{\lambda,\sigma}) = 0 & \text{in } \mathbb{R}^{n+1}_+, \\ \partial^\gamma_\nu W_{\lambda,\sigma} = -\kappa_\gamma \left(\lim\limits_{x_{n+1}\to 0+} x_{n+1}^{1-2\gamma} \dfrac{\partial W_{\lambda,\sigma}}{\partial x_{n+1}} \right) = (-\Delta)^\gamma w_{\lambda,\sigma} = w_{\lambda,\sigma}^{\frac{n+2\gamma}{n-2\gamma}} & \text{on } \mathbb{R}^n. \end{cases} \tag{1-15}$$

(In light of the equation that $W_{\lambda,\sigma}$ solves, we say that $W_{\lambda,\sigma}$ is γ*-harmonic*. Refer to [Caffarelli and Silvestre 2007]. For future use, let $W_\lambda = W_{\lambda,0}$ and $w_\lambda = w_{\lambda,0}$.) Moreover, if $S_{n,\gamma} > 0$ denotes the best constant one can achieve in (1-12) and $(\mathbb{S}^n, [g_c])$ is the standard unit n-dimensional sphere, then

$$\Lambda^\gamma(\mathbb{S}^n, [g_c]) = S_{n,\gamma}^{-1} \kappa_\gamma = \left(\int_{\mathbb{R}^n} w_{\lambda,\sigma}^{\frac{2n}{n-2\gamma}}\, d\bar{x} \right)^{\frac{2\gamma}{n}}. \tag{1-16}$$

Related to this fact, we have the following compactness result.

Proposition B. *Let $n > 2\gamma$, $\gamma \in (0,1)$ and (X^{n+1}, g^+) be an asymptotically hyperbolic manifold with the conformal infinity $(M^n, [\hat{h}])$. Also, assume that* (1-9) *is true. Then*

$$-\infty < \Lambda^\gamma(M, [\hat{h}]) \le \Lambda^\gamma(\mathbb{S}^n, [g_c]), \tag{1-17}$$

and the fractional Yamabe problem (1-7)–(1-8) *has a positive solution if the strict inequality holds.*

Refer to [González and Qing 2013, Sections 5 and 6] for its proof. Moreover since (1-17) automatically holds if the γ-Yamabe constant $\Lambda^\gamma(M, [\hat{h}])$ is negative or 0, *we assume that* $\Lambda^\gamma(M, [\hat{h}]) > 0$ *from now on.*

The purpose of this paper is to construct a proper nonzero test function $\Phi \in W^{1,2}(X, \rho^{1-2\gamma})$ such that $0 < \bar{I}^\gamma_{\hat{h}}[\Phi] < \Lambda^\gamma(\mathbb{S}^n, [g_c])$ when $\gamma \in (0,1)$, (X^{n+1}, g^+) is an asymptotically hyperbolic manifold, (1-9) holds and

- M^n has a point where the mean curvature H is negative, $n \ge 2$ and $\gamma \in \left(0, \frac{1}{2}\right)$; or
- M^n is the nonumbilic boundary of X^{n+1}, $n \ge 4$ and assumption (1-18) holds; or
- M^n is the umbilic boundary of X^{n+1}, the covariant derivative $R_{\rho\rho;\rho}[\bar{g}]$ of the Ricci tensor $R_{\rho\rho}[\bar{g}]$ on $(\overline{X}, \bar{g})$ is negative at a certain point of M, $n > 3 + 2\gamma$ and hypothesis (1-18) holds (where ρ is the geodesic defining function associated to $(M, \hat{h})$ and $\bar{g} = \rho^2 g^+$); or

- M^n is the umbilic but nonlocally conformally flat boundary of X^{n+1}, $n > 4+2\gamma$ and condition (1-19) is satisfied; or
- X^{n+1} is Poincaré–Einstein, either M^n is $n \geq 3$ and locally conformally flat or $n = 2$, the expansion (1-21) of the Green's function $G(\cdot, y)$ holds in a neighborhood of an arbitrarily chosen point $y \in M$ and the constant-order term A of $G(\cdot, y)$ is positive.

Once it is achieved, Proposition B will imply the existence of a positive solution to (1-2) automatically. The natural candidate for a positive test function is certainly the standard bubble, possibly truncated. Indeed, this is a good choice for the first and third cases mentioned above. Nevertheless, to cover lower-dimensional manifolds or locally conformally flat boundaries, it is necessary to find more accurate test functions than the truncated bubbles; cf. [González and Qing 2013; González and Wang 2017]. To take into account the second and fourth situations, we shall add a correction term on the bubble by adapting the idea of Marques [2007] and Almaraz [2010]. For the fifth case, we will construct an appropriate test function by utilizing the Green's function $G(\cdot, y)$. In the local situation $\gamma = 1$, such an approach was successfully applied by Schoen [1984]. His idea was later extended by Escobar [1992a] in the work on the boundary Yamabe problem, which has close relationship to the fractional Yamabe problem with $\gamma = \frac{1}{2}$, as discussed.

Let π be the second fundamental form of $(M, \hat{h}) \subset (\overline{X}, \bar{g})$. The boundary M is called *umbilic* if the tensor $T := \pi - H\bar{g}$ vanishes on M. Also M is *nonumbilic* if it possesses a point at which $T \neq 0$. Our first main result reads as follows:

Theorem 1.1. *Suppose that (X^{n+1}, g^+) is an asymptotically hyperbolic manifold, $(M, [\hat{h}])$ is its conformal infinity and* (1-9) *holds. Assume also that ρ is a geodesic defining function of $(M, \hat{h})$ and $\bar{g} = \rho^2 g^+ = d\rho^2 \oplus h_\rho$ near $M = \{\rho = 0\}$. If either*

- *$n \geq 2$, $\gamma \in \left(0, \frac{1}{2}\right)$ and M^n has a point at which the mean curvature H is negative; or*
- *$n \geq 4$, $\gamma \in (0, 1)$, M^n is the nonumbilic boundary of X^{n+1} and*

$$R[g^+] + n(n+1) = o(\rho^2) \quad \text{as } \rho \to 0 \text{ uniformly on } M, \tag{1-18}$$

then the γ-Yamabe problem is solvable — namely, (1-2) *has a positive solution.*

Remark 1.2. (1) As pointed out in [González and Qing 2013], we are only permitted to change the metric on the conformal infinity M. Once the boundary metric $\hat{h}$ is fixed, the geodesic boundary-defining function ρ and a compact metric $\bar{g}$ on X are automatically determined by the relations $|d\rho|_{\rho^2 g^+} = 1$ and $\bar{g} = \rho^2 g^+$. This is a huge difference between the fractional Yamabe problem (especially, with $\gamma = \frac{1}{2}$) and the boundary Yamabe problem, in that one has a freedom of conformal change of the metric in the whole manifold X when he/she is concerned with the boundary Yamabe problem.

Due to this reason, while it is possible to make the "extrinsic" metric H vanish at a point by a conformal change in the boundary Yamabe problem, one cannot do the same thing in the setting of the fractional Yamabe problem. This forced us to separate the cases in the statement of Theorem 1.1.

(2) As a particular consequence of the previous discussion, the Ricci tensor $R_{\rho\rho}[\bar{g}](y)$ of $(X, \bar{g})$ evaluated at a point y on M is governed by $\hat{h}$ and (1-18) (see Lemma 2.4). In the boundary Yamabe problem, Escobar [1992a] could choose a metric in X such that $R_{ij}[\hat{h}](y) = 0$ and $R_{\rho\rho}[\bar{g}](y) = 0$ simultaneously.

Moreover, by putting (1-6) and (1-18) together, we get

$$E(\rho) = \frac{n-2\gamma}{4n} R[\bar{g}]\,\rho^{1-2\gamma} + o(\rho^{1-2\gamma}) \quad \text{near } M.$$

Hence, on account of the energy expansion, (1-18) is the very condition that makes the boundary Yamabe problem and the $\frac{1}{2}$-Yamabe problem identical modulo the remainder. Refer to Subsections 2C and 2D.

(3) The sign of the mean curvature at a fixed point on M and (1-18) are "intrinsic" curvature conditions of an asymptotically hyperbolic manifold in the sense that these properties are independent of the choice of a representative of the class $[\hat{h}]$. Refer to Lemma 2.1 below for its proof. Also Lemma 2.3 claims that (1-18) implies $H = 0$ on M.

(4) Note also that $2 + 2\gamma \in \mathbb{N}$ and $\gamma \in (0, 1)$ if and only if $\gamma = \frac{1}{2}$, and the boundary Yamabe problem on nonumbilic manifolds in dimension $n = 2 + 2\gamma = 3$ was covered in [Marques 2007]. We expect that the strategy suggested in that paper can be applied for $\frac{1}{2}$-Yamabe problem in the same setting.

We next consider the case when the boundary M is umbilic but either $R_{\rho\rho;\rho}[\bar{g}] < 0$ at some point on M or it is nonlocally conformally flat.

Theorem 1.3. *Suppose that* (X^{n+1}, g^+) *is an asymptotic hyperbolic manifold such that* (1-9) *holds and the boundary* $(M^n, [\hat{h}])$ *is umbilic. If either*

- $n > 3 + 2\gamma,\ \gamma \in (0, 1)$, *that is, either* $n \geq 5$ *and* $\gamma \in (0, 1)$ *or* $n = 4$ *and* $\gamma \in \left(0, \frac{1}{2}\right)$, *the tensor* $R_{\rho\rho;\rho}[\bar{g}]$ *is negative at a certain point of* M *and* (1-18) *is valid*; *or*
- $n > 4 + 2\gamma,\ \gamma \in (0, 1)$, *that is, either* $n \geq 6$ *and* $\gamma \in (0, 1)$ *or* $n = 5$ *and* $\gamma \in \left(0, \frac{1}{2}\right)$, *there is a point* $y \in M$ *such that the Weyl tensor* $W[\hat{h}]$ *on* M *is nonzero at* y *and*

$$\begin{cases} R[g^+] + n(n+1) = o(\rho^4), \\ \partial_{\bar{x}}^m (R[g^+] + n(n+1)) = o(\rho^2) \quad (m = 1, 2), \\ \partial_{\rho}^m (R[g^+] + n(n+1)) = o(\rho^2) \quad (m = 1, 2) \end{cases} \tag{1-19}$$

  as $\rho \to 0$ *uniformly on* M,

then the γ*-Yamabe problem is solvable. Here* $\bar{x}$ *is a coordinate on* M.

Remark 1.4. (1) As we will see later, the main order of the energy for the fractional Yamabe problem (1-2) is ϵ^4 on an umbilic but nonlocally conformal flat boundary M, while it is ϵ^2 on a nonumbilic boundary; see (2-11), (2-14), (3-14) and (3-16). This explains why the necessary decay rate of $R[g^+] + n(n+1)$ to 0 as $\rho \to 0$ in Theorem 1.3 should be ρ^2-times as fast as that in Theorem 1.1.

On the other hand, (1-19) is responsible for determining all the values of quantities which emerge in the coefficient of ϵ^4 in the energy (such as $R_{,ii}[\bar{g}](y)$ and $R_{NN,ii}[\bar{g}](y)$ — see Lemma 3.2) and making the term $(n(n+1) + R[g^+])\rho^{-2}$ in $E(\rho)$ to be ignorable.

(2) Owing to Lemmas 2.1 and 2.3, condition (1-19) is again intrinsic and sufficient to deduce that $H = 0$ on M. Moreover every Poincaré–Einstein manifold satisfies (1-19).

In [González and Wang 2017, Lemma 2.3], it is proved that the sign of the tensor $R_{\rho\rho;\rho}[\bar{g}]$ at a fixed point on M is intrinsic.

(3) It is notable that $4 + 2\gamma \in \mathbb{N}$ and $\gamma \in (0, 1)$ if and only if $\gamma = \frac{1}{2}$, and the boundary Yamabe problem for $n = 4 + 2\gamma = 5$ was studied in [Almaraz 2010]. We believe that Theorem 1.3 can be extended to the case $\gamma = \frac{1}{2}$, $n = 5$, $W[\hat{h}] \neq 0$ on M and (1-19) is valid.

In order to describe the last result, we first have to take into account of the existence of a Green's function under our setting.

Proposition 1.5. *Suppose that all the hypotheses of Theorem A hold true, including* (1-9), *and* $H = 0$ *on* M. *In addition, assume further that* $\Lambda^\gamma(M, [\hat{h}]) > 0$. *Then for each* $y \in M$, *there exists a Green's function* $G(x, y)$ *on* $\overline{X} \setminus \{y\}$ *which satisfies*

$$\begin{cases} -\operatorname{div}_{\bar{g}}(\rho^{1-2\gamma} \nabla G(\cdot\,, y)) + E(\rho)\, G(\cdot\,, y) = 0 & \text{in } (X, \bar{g}), \\ \partial_\nu^\gamma G(\cdot\,, y) = \delta_y & \text{on } (M, \hat{h}) \end{cases} \tag{1-20}$$

in the distribution sense, where δ_y *is the Dirac measure at* y. *The function* G *is unique and positive on* $\overline{X}$.

The proof is postponed until Section 4A. The readers may compare the above result with [Guillarmou and Qing 2010]. Based on standard elliptic regularity and the facts that if $(X, \bar{g})$ is the Poincaré half-plane $(\mathbb{R}^{n+1}_+, x_{n+1}^{-2}\, dx)$, then

$$G(x, \bar{y}) = \frac{g_{n,\gamma}}{|(\bar{x} - \bar{y}, x_{n+1})|^{n-2\gamma}} \quad \text{for all } (\bar{x}, x_{n+1}) \in \mathbb{R}^{n+1}_+ \text{ and } \bar{y} \in \mathbb{R}^n,$$

and that the compactified metric $\bar{g}$ on $\overline{X}$ of a Poincaré–Einstein manifold (X, g^+) can be assumed to be Euclidean up to order $|x|^n$ in its coordinate $x \in \mathbb{R}^{n+1}_+$ (refer to Lemma 4.3 below), we expect the following.

Conjecture 1.6 (expansion of the Green's function). Assume that $\gamma \in (0, 1)$, $n > 2\gamma$ and (X^{n+1}, g^+) is Poincaré–Einstein. Also, suppose that $\Lambda^\gamma(M, [\hat{h}]) > 0$ and that either $(M^n, [\hat{h}])$ has $n \geq 3$ and is locally conformally flat or $n = 2$. Fix any $y \in M$. Then there exists a local coordinate x of the compact manifold $(\overline{X}, \bar{g})$ around y (identified with $0 \in \mathbb{R}^n$) defined in a small closed neighborhood $\mathcal{N} \subset \overline{\mathbb{R}}^{n+1}_+$ of 0 such that

$$G(x, 0) = g_{n,\gamma} |x|^{-(n-2\gamma)} + A + \Psi(x) \quad \text{for } x \in \mathcal{N}. \tag{1-21}$$

Here $g_{n,\gamma} > 0$ is a number that appeared in (1-13), $A \in \mathbb{R}$ and Ψ is a function in $\mathcal{N}$ satisfying

$$|\Psi(x)| \leq C|x|^{\min\{1, 2\gamma\}} \quad \text{and} \quad |\nabla\Psi(x)| \leq C|x|^{\min\{0, 2\gamma-1\}} \quad \text{for } x \in \mathcal{N} \tag{1-22}$$

for some constant $C > 0$.

Now we can state our third main theorem.

Theorem 1.7. *Suppose that $\gamma \in (0,1)$, $n > 2\gamma$ and (X^{n+1}, g^+) is a Poincaré–Einstein manifold with conformal infinity $(M^n, [\hat{h}])$. Let ρ be a geodesic defining function for $(M, \hat{h})$ and $\bar{g} = \rho^2 g^+$. If* (1-9) *holds, Conjecture 1.6 is valid, $A > 0$, and either M^n has $n \geq 3$ and is locally conformally flat or $n = 2$, then the fractional Yamabe problem is solvable.*

Remark 1.8. (1) Let us set a 2-tensor

$$F = \rho(\mathrm{Ric}[g^+] + ng^+) \quad \text{in } X,$$

which is identically 0 if (X, g^+) is Poincaré–Einstein. As a matter of fact, if M is locally conformally flat, the only property of the tensor F necessary to derive Theorem 1.7 is that $\partial_\rho^m F|_{\rho=0} = 0$ for $m = 0, \dots, n-1$ (refer to Lemma 4.3). We guess that (1-21) and (1-22) are still valid under this assumption. Similarly, for the case $n = 2$, the assumption $\partial_\rho^m F|_{\rho=0} = 0$ for $m = 0, 1$ would suffice.

(2) Since (X^{n+1}, g^+) is Poincaré–Einstein, the second fundamental form on M is trivial. Particularly, the mean curvature H on M vanishes and M is umbilic.

(3) Suppose that we are in the local case $\gamma = 1$, and either $n \geq 7$ or M is locally conformally flat. Then, as shown in [Lee and Parker 1987, Lemma 6.4], the expansion (1-21) is valid. Furthermore, the classical positive mass theorem of Schoen and Yau [1979a; 1979b; 1988] states that $A \geq 0$, and the positivity condition $A > 0$ holds if and only if $(M, \hat{h})$ is not conformally diffeomorphic to the standard sphere $\mathbb{S}^n$. Determining the sign of A at each point $y \in M$ is a still natural problem for $\gamma \in (0,1)$. However, it is difficult to perform, because A may be a nonlocal quantity, namely, one depending on the whole geometry of (X, g^+) and $(M, [\hat{h}])$.

This paper is organized as follows: In Section 2, we establish Theorem 1.1 by intensifying the ideas of Marques [2007] and González and Qing [2013]. Section 3 provides the proof of Theorem 1.3, which further develops the approach of Almaraz [2010] and González and Wang [2017]. In Section 4, Theorem 1.7 is achieved, which can be understood as a sort of generalization of the results of Schoen [1984] and Escobar [1992a]. In particular, Section 4A is devoted to investigating the existence and some qualitative properties of a Green's function (i.e., Proposition 1.5). Then we are concerned with the case that M is locally conformally flat (in Section 4B) and 2-dimensional (in Section 4C). Finally, we examine the asymptotic behavior of the bubble $W_{1,0}$ near infinity in Appendix A, and compute some integrations regarding $W_{1,0}$, which are needed in the energy expansions in Appendix B.

Notation. • The Einstein convention is used throughout the paper. The indices i, j, k and l always take values from 1 to n, and a and b range over values from 1 to $n+1$.

• For a tensor T, notations $T_{;a}$ and $T_{,a}$ indicate covariant differentiation and partial differentiation of T, respectively.

• For a tensor T and a number $q \in \mathbb{N}$, we use

$$\mathrm{Sym}_{i_1 \cdots i_q} T_{i_1 \cdots i_q} = \frac{1}{q!} \sum_{\sigma \in S_q} T_{i_{\sigma(1)} \cdots i_{\sigma(q)}},$$

where S_q is the group of all permutations of q elements.

• We let $N = n+1$. Also, for $x \in \mathbb{R}^N_+ := \{(x_1, \dots, x_n, x_N) \in \mathbb{R}^N : x_N > 0\}$, we write $\bar{x} = (x_1, \dots, x_n, 0) \in \partial\mathbb{R}^N_+ \simeq \mathbb{R}^n$ and $r = |\bar{x}|$.

• For $n > 2\gamma$, we set $p = (n+2\gamma)/(n-2\gamma)$.

• For any $\varrho > 0$, let $B^n(0, \varrho)$ and $B^N_+(0, \varrho)$ be the n-dimensional ball and the N-dimensional upper half-ball centered at 0 whose radius is ϱ, respectively.

• $|\mathbb{S}^{n-1}|$ is the surface area of the $(n-1)$-dimensional unit sphere $\mathbb{S}^{n-1}$.

• For any $t \in \mathbb{R}$, let $t_+ = \max\{0, t\} \geq 0$ and $t_- = \max\{0, -t\} \geq 0$ so that $t = t_+ - t_-$.

• For $\gamma \in (0, 1)$, the space $H^\gamma(M)$ is the completion of $C^\infty(M)$ with respect to the norm which one obtains by pulling back

$$u \in C^\infty_c(\mathbb{R}^n) \mapsto \left(\int_{\mathbb{R}^n} u^2 \, d\bar{x} + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(\bar{x}) - u(\bar{y})|^2}{|\bar{x} - \bar{y}|^{n+2\gamma}} \, d\bar{x} \, d\bar{y} \right)^{\frac{1}{2}}$$

to M through coordinate charts.

• The space $D^{1,2}(\mathbb{R}^N_+, x_N^{1-2\gamma})$ denotes the completion of $C^\infty_c(\overline{\mathbb{R}}^N_+)$ with respect to the norm

$$U \mapsto \left(\int_{\mathbb{R}^N_+} x_N^{1-2\gamma} |\nabla U|^2 \, dx \right)^{\frac{1}{2}},$$

and the space $W^{1,2}(X, \rho^{1-2\gamma})$ denotes the completion of $C^\infty_c(\overline{X})$ with respect to the norm

$$U \mapsto \left(\int_X \rho^{1-2\gamma} (|\nabla U|^2_{\bar{g}} + U^2) \, dv_{\bar{g}} \right)^{\frac{1}{2}}.$$

In light of Theorem A, $W^{1,2}(X, \rho^{1-2\gamma})$ is the natural functional space for the fractional Yamabe problem.

• The following positive constants are given in (1-8), (1-13) and (1-14):

$$\begin{aligned} \kappa_\gamma &= \frac{\Gamma(\gamma)}{2^{1-2\gamma}\Gamma(1-\gamma)}, \quad p_{n,\gamma} = \frac{\Gamma\left(\frac{n+2\gamma}{2}\right)}{\pi^{\frac{n}{2}}\Gamma(\gamma)}, \\ g_{n,\gamma} &= \frac{\Gamma\left(\frac{n-2\gamma}{2}\right)}{\pi^{\frac{n}{2}} 2^{2\gamma}\Gamma(\gamma)}, \qquad \alpha_{n,\gamma} = 2^{\frac{n-2\gamma}{2}} \left(\frac{\Gamma\left(\frac{n+2\gamma}{2}\right)}{\Gamma\frac{1}{2}(n-2\gamma)} \right)^{\frac{n-2\gamma}{4\gamma}}. \end{aligned} \tag{1-23}$$

• $C > 0$ is a generic constant which may vary from line to line.

2. Nonminimal and nonumbilic conformal infinities

2A. ***Geometric background.*** We begin this section by proving that the sign of the mean curvature, (1-18) and nonumbilicity of a point on M are intrinsic conditions.

Lemma 2.1. *Suppose that (X, g^+) is an asymptotically hyperbolic manifold with conformal infinity $(M, [\hat{h}])$. Moreover, let ρ and $\tilde{\rho}$ be the geodesic boundary-defining functions associated to two representatives $\hat{h}$ and $\tilde{h}$ of the class $[\hat{h}]$, respectively. We also define $\bar{g} = \rho^2 g^+$ and $\tilde{g} := \tilde{\rho}^2 g^+$, denote by*

$\pi = -\bar{g}_{,N}/2$ *and* $\tilde{\pi}$ *the second fundamental forms of* $(M, \hat{h}) \subset (\overline{X}, \bar{g})$ *and* $(M, \tilde{h}) \subset (\overline{X}, \tilde{g})$, *respectively, and set* $H = \bar{g}^{ij}\pi_{ij}/n$ *and* $\tilde{H} = \tilde{g}^{ij}\tilde{\pi}_{ij}/n$. *Then we have*

$$C^{-1} \le \frac{\tilde{\rho}}{\rho} \le C \quad \text{in } X \qquad \text{and} \qquad H = \frac{\tilde{\rho}}{\rho}\bigg|_{\rho=0} \tilde{H} \quad \text{on } M \tag{2-1}$$

for some $C > 1$. *Furthermore if* $H = 0$ *on* M, *then*

$$\pi = \frac{\rho}{\tilde{\rho}}\bigg|_{\rho=0} \tilde{\pi} \quad \text{on } M. \tag{2-2}$$

Proof. The assertion on H in (2-1) is proved in [González and Qing 2013, Lemma 2.3]. For the first inequality in (2-1), it suffices to observe that $\tilde{\rho}/\rho$ is bounded above and bounded away from 0 near M. Indeed, this follows from the fact that

$$\tilde{h} = \tilde{g}|_M = \tilde{\rho}^2 g^+|_M = \left(\frac{\tilde{\rho}}{\rho}\right)^2 \bar{g}|_M = \left(\frac{\tilde{\rho}}{\rho}\right)^2 \hat{h} \quad \text{on } M.$$

Let us define tensors $T = \pi - H\bar{g}$ and $\widetilde{T} = \tilde{\pi} - \tilde{H}\tilde{g}$ on M. Then we see from [Escobar 1992b, Proposition 1.2] that

$$\tilde{\pi} = \widetilde{T} = \frac{\tilde{\rho}}{\rho} T = \frac{\tilde{\rho}}{\rho}\pi \quad \text{on } M$$

provided $H = 0$ on M, which confirms (2-2). □

Given any fixed point $y \in M$, let $\bar{x} = (x_1, \dots, x_n)$ be normal coordinates on M at y (identified with 0) and $x_N = \rho$. In other words, let $x = (\bar{x}, x_N)$ be *Fermi coordinates*. The following lemma provides the expansion of the metric $\bar{g}$ near $y = 0$. See [Escobar 1992a, Lemma 3.1] for its proof.

Lemma 2.2. *Let* $(\overline{X}, \bar{g})$ *be a compact manifold with boundary* $(M, \hat{h})$ *and* $y \in M$. *Then, in terms of Fermi coordinates around* y, *it holds that*

$$\sqrt{|\bar{g}|}(x) = 1 - nHx_N + \tfrac{1}{2}\big(n^2H^2 - \|\pi\|^2 - R_{NN}[\bar{g}]\big)x_N^2 - H_{,i}x_ix_N - \tfrac{1}{6}R_{ij}[\hat{h}]x_ix_j + O(|x|^3)$$

and

$$\bar{g}^{ij}(x) = \delta_{ij} + 2\pi_{ij}x_N + \tfrac{1}{3}R_{ikjl}[\hat{h}]x_kx_l + \bar{g}^{ij}_{\ ,Nk}x_Nx_k + \big(3\pi_{ik}\pi_{kj} + R_{iNjN}[\bar{g}]\big)x_N^2 + O(|x|^3)$$

near y (*identified with a small half-ball* $B^N_+(0, 2\eta_0)$ *near* 0 *in* $\mathbb{R}^N_+$). *Here* $\|\pi\|^2 = \hat{h}^{ik}\hat{h}^{jl}\pi_{ij}\pi_{kl}$ *is the square of the norm of the second fundamental form* π *on* $(M, \hat{h}) \subset (\overline{X}, \bar{g})$, $R_{ikjl}[\hat{h}]$ *is a component of the Riemannian curvature tensor on* M, $R_{iNjN}[\bar{g}]$ *is that of the Riemannian curvature tensor in* X, $R_{ij}[\hat{h}] = R_{ikjk}[\hat{h}]$ *and* $R_{NN}[\bar{g}] = R_{iNiN}[\bar{g}]$. *Every tensor in the expansions is computed at* $y = 0$.

Now notice that the transformation law of the scalar curvature [Escobar 1992a, (1.1)] implies

$$R[g^+] + n(n+1) = 2n\frac{\partial_\rho\sqrt{|\bar{g}|}}{\sqrt{|\bar{g}|}}\rho + R[\bar{g}]\rho^2. \tag{2-3}$$

It readily shows that (1-18) and (1-19) indicate $H = 0$ on M.

Lemma 2.3. *Suppose that (X, g^+) is an asymptotically hyperbolic manifold with conformal infinity $(M, [\hat{h}])$. If $R[g^+] + n(n+1) = o(\rho)$ as $\rho \to 0$, then $H = 0$ on M.*

Proof. Fix any $y \in M$. By (2-3), we have

$$o(1) = 2n \frac{\partial_\rho \sqrt{|\bar{g}|}(y)}{\sqrt{|\bar{g}|}(y)} + R[\bar{g}](y)\rho + o(1) = -2n^2 H(y) + o(1)$$

as a point tends to y. This implies $H(y) = 0$, and therefore the assertion follows. □

We next select a good background metric on X under the validity of hypothesis (1-18).

Lemma 2.4. *Let (X, g^+) be an asymptotically hyperbolic manifold such that condition* (1-18) *holds. Then the conformal infinity $(M, [\hat{h}])$ admits a representative $\hat{h} \in [\hat{h}]$, the geodesic boundary-defining function ρ and the metric $\bar{g} = \rho^2 g^+$ satisfying*

$$H = 0 \quad \text{on } M, \qquad R_{ij}[\hat{h}](y) = 0 \qquad \text{and} \qquad R_{\rho\rho}[\bar{g}](y) = \frac{1-2n}{2(n-1)} \|\pi(y)\|^2 \tag{2-4}$$

for a fixed point $y \in M$.

Proof. According to [Lee and Parker 1987, Theorem 5.2], one may choose a representative $\hat{h}$ of the conformal class $[\hat{h}]$ such that $R_{ij}[\hat{h}](y) = 0$. Additionally Lemmas 2.3 and 2.1 ensure that $H = 0$ on M for any $\hat{h} \in [\hat{h}]$. Hence assumption (1-18) can be interpreted as

$$\begin{aligned} o(1) &= 2n \frac{\partial_\rho \sqrt{|\bar{g}|}}{\rho \sqrt{|\bar{g}|}} + R[\bar{g}] = \frac{n}{\rho} \bar{g}^{ab} \bar{g}_{ab,\rho} + R[\bar{g}] = n(\bar{g}^{ab}{}_{,\rho}\, \bar{g}_{ab,\rho} + \bar{g}^{ab} \bar{g}_{ab,\rho\rho}) + R[\bar{g}] + o(1) \\ &= -2n(R_{\rho\rho}[\bar{g}] + \|\pi\|^2) + \big(2R_{\rho\rho}[\bar{g}] + \|\pi\|^2 + R[\hat{h}] - H^2\big) + o(1) \end{aligned}$$

as $\rho \to 0$, where we used $H = 0$ on M for the third equality and the Gauss–Codazzi equation for the fourth equality; see the proof of Lemmas 3.1 and 3.2 of [Escobar 1992a]. Taking the limit to $y \in M$, we get

$$0 = 2(1-n) R_{\rho\rho}[\bar{g}](y) + (1-2n)\|\pi(y)\|^2.$$

The third equality of (2-4) is its direct consequence. □

Lastly, we recall the function E in (1-5) and (1-6). In a collar neighborhood of M where $\rho = x_N$, it can be seen that

$$E(x_N) = \frac{n-2\gamma}{4n} \big[R[\bar{g}] - (n(n+1) + R[g^+]) x_N^{-2}\big] x_N^{1-2\gamma} = -\tfrac{1}{2}(n-2\gamma) \frac{\partial_N \sqrt{|\bar{g}|}}{\sqrt{|\bar{g}|}} x_N^{-2\gamma}, \tag{2-5}$$

where the second equality holds because of (2-3).

2B. ***Nonminimal conformal infinity.*** Let $y \in M$ be a point identified with $0 \in \mathbb{R}^n$ such that $H(y) < 0$ and $B_+^N(0, 2\eta_0) \subset \mathbb{R}_+^N$ be its neighborhood which appeared in Lemma 2.2. Also, we select any smooth radial cut-off function $\psi \in C_c^\infty(\mathbb{R}_+^N)$ such that $\psi = 1$ in $B_+^N(0, \eta_0)$ and 0 in $\mathbb{R}_+^N \setminus B_+^N(0, 2\eta_0)$. In this subsection, we shall show that $\bar{I}_{\hat{h}}^\gamma[\psi W_\epsilon] < \Lambda^\gamma(\mathbb{S}^n, [g_c])$ for any $n \geq 2$ and $\gamma \in \big(0, \frac{1}{2}\big)$, where $W_\epsilon = W_{\epsilon,0}$ as before.

Before starting the computation, let us make one useful observation: Assume that $n > m + 2\gamma$ for a certain $m \in \mathbb{N}$. Then we get from (A-3) and (A-4) that

$$\int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |x|^{m+1} |\nabla W_\epsilon|^2 \, dx = \eta_0^{1-\zeta} \int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |x|^{m+\zeta} |\nabla W_\epsilon|^2 \, dx = O(\epsilon^{m+\zeta}) = o(\epsilon^m) \quad (2\text{-}6)$$

by choosing a small number $\zeta > 0$ such that $n > m + 2\gamma + \zeta$.

Proposition 2.5. *Suppose that (X^{n+1}, g^+) is an asymptotically hyperbolic manifold with conformal infinity $(M, [\hat{h}])$ and $y \in M$ is a point such that $H(y) < 0$. Then for any $\epsilon > 0$ small, $n \geq 2$ and $\gamma \in \left(0, \frac{1}{2}\right)$, we have*

$$\bar{I}_{\hat{h}}^\gamma[\psi W_\epsilon] \leq \Lambda^\gamma(\mathbb{S}^n, [g_c]) + \epsilon \underbrace{\frac{2n^2-2n+1-4\gamma^2}{2(1-2\gamma)} \, \frac{\kappa_\gamma \int_{\mathbb{R}_+^N} x_N^{2-2\gamma} |\nabla W_1|^2 \, dx}{\left(\int_{\mathbb{R}^n} w_1^{p+1} \, dx\right)^{\frac{n-2\gamma}{n}}}}_{>0} H(y) + o(\epsilon) < \Lambda^\gamma(\mathbb{S}^n, [g_c]), \quad (2\text{-}7)$$

where $\bar{I}_{\hat{h}}^\gamma$ is the γ-Yamabe functional given in (1-11), *and $\Lambda^\gamma(\mathbb{S}^n, [g_c])$ and κ_γ are positive constants introduced in* (1-16) *and* (1-23).

Proof. Since the proof is essentially the same as that of [Choi and Kim 2017, Proposition 6.1], we briefly sketch it. By Lemma 2.2 and (2-6), we discover

$$\begin{aligned} &\int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_\epsilon|_{\bar{g}}^2 \, dv_{\bar{g}} \\ &\quad = \int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_1|^2 \, dx + \epsilon H \left(2 \int_{\mathbb{R}_+^N} x_N^{2-2\gamma} |\nabla_{\bar{x}} W_1|^2 \, dx - n \int_{\mathbb{R}_+^N} x_N^{2-2\gamma} |\nabla W_1|^2 \, dx \right) + o(\epsilon) \end{aligned}$$

and

$$\int_M (\psi W_\epsilon)^{p+1} \, dv_{\hat{h}} = \int_{B^n(0,\eta_0)} w_\epsilon^{p+1} (1 + O(|\bar{x}|^2)) \, d\bar{x} + O(\epsilon^n) = \int_{\mathbb{R}^n} w_1^{p+1} \, dx + o(\epsilon).$$

Moreover, according to Lemma 2.2 and (2-5), we have

$$\int_{B_+^N(0,\eta_0)} E(x_N) W_\epsilon^2 \, dv_{\bar{g}} = \tfrac{1}{2} n(n-2\gamma) \epsilon H \int_{\mathbb{R}_+^N} x_N^{-2\gamma} W_1^2 \, dx + o(\epsilon).$$

Thus the above estimates and Lemma B.3 confirm (2-7). □

Unlike the other existence results to be discussed later, we need to assume that $\gamma \in \left(0, \frac{1}{2}\right)$ for Proposition 2.5. Such a restriction is necessary in two reasons: First of all, $\gamma \in \left(0, \frac{1}{2}\right)$ is necessary for the function $x_N^{-2\gamma} W_1^2$ to be integrable in $\mathbb{R}_+^N$. Secondly the mean curvature H should vanish for $\gamma \in \left(\frac{1}{2}, 1\right)$ to guarantee the validity of the extension theorem (Theorem A).

2C. ***Nonumbilic conformal infinity: higher-dimensional cases.*** We fix a nonumbilic point $y = 0 \in M$. Let also $B_+^N(0, 2\eta_0) \subset \mathbb{R}_+^N$ be a small neighborhood of 0 and $\psi \in C_c^\infty(B_+^N(0, 2\eta_0))$ a cut-off function chosen in the previous subsection.

Lemma 2.6. *Let $J^{\gamma}_{\hat{h}}$ be the energy functional defined as*

$$J^{\gamma}_{\hat{h}}[U;X]=\int_X (\rho^{1-2\gamma}|\nabla U|^2_{\bar{g}}+E(\rho)U^2)\,dv_{\bar{g}}\quad \text{for any } U\in W^{1,2}(X,\rho^{1-2\gamma}). \tag{2-8}$$

Assume also that (2-4) *holds. Then for any $\epsilon>0$ small, $n>2+2\gamma$ and $\gamma\in(0,1)$, it is valid that*

$$\begin{aligned}J^{\gamma}_{\hat{h}}[\psi W_\epsilon;B^N_+(0,\eta_0)]&=\int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma}|\nabla W_1|^2\,dx\\&\quad+\epsilon^2\|\pi\|^2\big[-\tfrac{1}{2}(1+b)\mathcal{F}_2+\tfrac{1}{n}(3+b)\mathcal{F}_3+\tfrac{1}{2}(n-2\gamma)(1+b)\mathcal{F}_1\big]+o(\epsilon^2),\end{aligned} \tag{2-9}$$

where $b:=(1-2n)/(2n-2)$, $\|\pi\|$ is the norm of the second fundamental form at $y=0\in M$, and the values $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ are given in Lemma B.4.

Proof. We borrow the argument presented in [González and Qing 2013, Theorem 1.5]. According to Lemma 2.2 and (2-4), it holds that

$$\sqrt{|\bar{g}|}(\bar{x},x_N)=1-\tfrac{1}{2}(1+b)\|\pi\|^2x_N^2+O(|(\bar{x},x_N)|^3)\quad \text{in } B^N_+(0,\eta_0). \tag{2-10}$$

Hence we obtain from (2-6) that

$$\begin{aligned}&\int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma}|\nabla W_\epsilon|^2_{\bar{g}}\,dv_{\bar{g}}\\&\quad=\int_{\mathbb{R}^N_+} x_N^{1-2\gamma}|\nabla W_\epsilon|^2\,dx+\epsilon^2\bigg[(3\pi_{ik}\pi_{kj}+R_{iNjN}[\bar{g}])\int_{\mathbb{R}^N_+} x_N^{3-2\gamma}\partial_i W_1\partial_j W_1\,dx\\&\qquad\qquad-\tfrac{1}{2}(1+b)\|\pi\|^2\int_{\mathbb{R}^N_+} x_N^{3-2\gamma}|\nabla W_1|^2\,dx\bigg]+o(\epsilon^2).\end{aligned}$$

Also, by means of (2-5) and (2-10),

$$E(x_N)=\tfrac{1}{2}(n-2\gamma)(1+b)\|\pi\|^2x_N^{1-2\gamma}+O(|x|^2x_N^{-2\gamma})$$

for $x_N\ge 0$ small, so

$$\int_{B^N_+(0,\eta_0)} E(x_N)W_\epsilon^2\,dv_{\bar{g}}=\epsilon^2\tfrac{1}{2}(n-2\gamma)(1+b)\|\pi\|^2\int_{\mathbb{R}^N_+} x_N^{1-2\gamma}W_1^2\,dx+o(\epsilon^2).$$

Collecting every calculation, we discover (2-9). □

The previous lemma ensures the existence of a positive solution to (1-2) for nonumbilic conformal infinity M^n with $n\in\mathbb{N}$ sufficiently high.

Corollary 2.7. *Assume that (X^{n+1},g^+) is an asymptotically hyperbolic manifold and $\hat{h}$ is the representative of the conformal infinity M found in Lemma 2.2. If $n>2+2\gamma$ and $\gamma\in(0,1)$, we have*

$$\bar{I}^{\gamma}_{\hat{h}}[\psi W_\epsilon]\le \Lambda^{\gamma}(\mathbb{S}^n,[g_c])-\epsilon^2\mathcal{C}'(n,\gamma)\,\Lambda^{\gamma}(\mathbb{S}^n,[g_c])^{-\frac{n-2\gamma}{2\gamma}}\kappa_\gamma\,|\mathbb{S}^{n-1}|A_3B_2\|\pi\|^2+o(\epsilon^2), \tag{2-11}$$

where the positive constants $\Lambda^\gamma(\mathbb{S}^n,[g_c])$, κ_γ, A_3 *and* B_2 *are introduced in* (1-16), (1-23) *and* (B-3), *respectively*, *and* $\mathcal{C}'(n,\gamma)$ *is the number given by*

$$\mathcal{C}'(n,\gamma)=\frac{3n^2+n(16\gamma^2-22)+20(1-\gamma^2)}{8n(n-1)(1-\gamma^2)}. \tag{2-12}$$

Proof. Estimate (2-11) comes from Lemmas 2.6 and B.4 and the computations made in the proof of [González and Qing 2013, Theorem 1.5]. The details are left to the reader. □

By (2-2), we still have that $\pi\neq 0$ at $y\in M$, even after picking a new representative of the conformal infinity. Furthermore, the number $\mathcal{C}'(n,\gamma)$ is positive when $n\geq 4$ for $\gamma>\sqrt{\frac{5}{11}}\simeq 0.674$, $n\geq 5$ for $\gamma>\frac{1}{2}$, $n\geq 6$ for $\gamma>\sqrt{\frac{1}{19}}\simeq 0.229$ and $n\geq 7$ for any $\gamma>0$. Hence, in this regime, one is able to deduce the existence of a positive solution of (1-2) by testing the truncated standard bubble into the γ-Yamabe functional.

2D. ***Nonumbilic conformal infinity: lower-dimensional cases.*** We recall the nonumbilic point $y\in M$ identified with the origin of $\mathbb{R}^N_+$, the small number $\eta_0>0$ and the cut-off function $\psi\in C_c^\infty(\mathbb{R}^N_+)$. Furthermore, we introduce

$$\Psi_\epsilon(\bar{x},x_N)=M_1\pi_{ij}x_ix_jx_Nr^{-1}\partial_rW_\epsilon=\epsilon\cdot\epsilon^{-\frac{n-2\gamma}{2}}\Psi_1(\epsilon^{-1}\bar{x},\epsilon^{-1}x_N) \tag{2-13}$$

for each $\epsilon>0$, where $M_1\in\mathbb{R}$ is a number to be determined later, the π_{ij} are the coefficients of the second fundamental form at y and $r=|\bar{x}|$. Our ansatz to deal with lower-dimensional cases is defined by

$$\Phi_\epsilon:=\psi(W_\epsilon+\Psi_\epsilon)\quad\text{in } X.$$

The definition of Φ_ϵ is inspired by [Marques 2007].

The main objective of this subsection is to prove:

Proposition 2.8. *Suppose that* (X^{n+1},g^+) *is an asymptotically hyperbolic manifold and* $\hat{h}$ *is the representative of the conformal infinity* M *satisfying* (2-4). *If* $n>2+2\gamma$ *and* $\gamma\in(0,1)$, *we have*

$$\bar{I}^\gamma_{\hat{h}}[\Phi_\epsilon]\leq\Lambda^\gamma(\mathbb{S}^n,[g_c])-\epsilon^2\mathcal{C}(n,\gamma)\,\Lambda^\gamma(\mathbb{S}^n,[g_c])^{-\frac{n-2\gamma}{2\gamma}}\kappa_\gamma\,|\mathbb{S}^{n-1}|A_3B_2\|\pi\|^2+o(\epsilon^2), \tag{2-14}$$

where $\mathcal{C}(n,\gamma)$ *is the number defined by*

$$\mathcal{C}(n,\gamma)=\frac{3n^2+n(16\gamma^2-22)+20(1-\gamma^2)}{8n(n-1)(1-\gamma^2)}+\frac{16(n-1)(1-\gamma^2)}{n(3n^2+n(2-8\gamma^2)+4\gamma^2-4)}.$$

It can be checked that $\mathcal{C}(n,\gamma)>0$ whenever $n\geq 4$ and $\gamma\in(0,1)$. Thus the above proposition, along with Proposition 2.5, justifies the statement of Theorem 1.1. We have $\mathcal{C}(3,\gamma)>0$ for $\gamma>\frac{1}{2}$, but it also holds that $n>2+2\gamma>3$. Therefore we get no result for $n=3$.

Proof of Proposition 2.8. The proof consists of three steps.

Step 1: energy in the half-ball $B^N_+(0,\eta_0)$. Since $\psi=1$ in $B^N_+(0,\eta_0)$, we discover

$$\begin{aligned}&J^\gamma_{\hat{h}}[\psi(W_\epsilon+\Psi_\epsilon);B^N_+(0,\eta_0)]\\&\quad=J^\gamma_{\hat{h}}[\psi W_\epsilon;B^N_+(0,\eta_0)]+2\int_{B^N_+(0,\eta_0)}x_N^{1-2\gamma}\langle\nabla W_\epsilon,\nabla\Psi_\epsilon\rangle_{\bar{g}}\,dv_{\bar{g}}+\int_{\mathbb{R}^N_+}x_N^{1-2\gamma}|\nabla\Psi_\epsilon|^2\,dx+o(\epsilon^2),\end{aligned} \tag{2-15}$$

where the functional $J^{\gamma}_{\hat{h}}$ is defined in (2-8). Moreover, we note from Lemma 2.2 that the mean curvature $H = \pi_{ii}/n$ vanishes at the origin, which yields

$$\begin{aligned}\int_{B^N_+(0,\eta_0)} & x_N^{1-2\gamma} \nabla W_\epsilon \cdot \nabla \Psi_\epsilon \, dx \\ &= \epsilon M_1 \int_{B^N_+(0,\eta_0/\epsilon)} x_N^{2-2\gamma} \pi_{ij} x_i x_j \big[2r^{-2} (\partial_r W_1)^2 + r \partial_r (r^{-1} \partial_r W_1) \big] \, dx \\ &\qquad + \epsilon M_1 \int_{B^N_+(0,\eta_0/\epsilon)} x_N^{1-2\gamma} \pi_{ij} x_i x_j r^{-1} (\partial_N W_1) \big[(\partial_r W_1) + x_N (\partial_{Nr} W_1) \big] \, dx = 0. \qquad (2\text{-}16)\end{aligned}$$

Hence we obtain from the definition (2-13) of Ψ_ϵ and (2-16) that

$$\begin{aligned} 2 \int_{B^N_+(0,\eta_0)} & x_N^{1-2\gamma} \langle \nabla W_\epsilon, \nabla \Psi_\epsilon \rangle_{\bar{g}} \, dv_{\bar{g}} \\ &= 2 \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma} \nabla W_\epsilon \cdot \nabla \Psi_\epsilon \, dx + 4\pi_{ij} \int_{\mathbb{R}^N_+} x_N^{2-2\gamma} \partial_i W_\epsilon \, \partial_j \Psi_\epsilon \, dx + o(\epsilon^2) \\ &= \epsilon^2 4 M_1 \pi_{ij} \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} x_i \big[2\pi_{jk} x_k r^{-2} (\partial_r W_1)^2 + \pi_{kl} x_k x_l x_j r^{-2} (\partial_r W_1) \, \partial_r (r^{-1} \partial_r W_1) \big] \, dx + o(\epsilon^2) \\ &= \epsilon^2 4 M_1 \left[\frac{2}{n} \mathcal{F}_3 + \frac{2}{n(n+2)} (-\mathcal{F}_3 + \mathcal{F}_4) \right] \|\pi\|^2 + o(\epsilon^2) \\ &= \epsilon^2 \left(\frac{4}{n} \right) M_1 |\mathbb{S}^{n-1}| A_3 B_2 \|\pi\|^2 + o(\epsilon^2), \end{aligned} \qquad (2\text{-}17)$$

where the constants $\mathcal{F}_3, \mathcal{F}_4$ as well as $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_5, \dots, \mathcal{F}_8$, are defined in Lemma B.4. In a similar fashion, it can be found that

$$\begin{aligned}\int_{\mathbb{R}^N_+} x_N^{1-2\gamma} |\nabla \Psi_\epsilon|^2 \, dx &= \epsilon^2 \frac{2M_1^2}{n(n+2)} (\mathcal{F}_3 - 2\mathcal{F}_4 + \mathcal{F}_5 + \mathcal{F}_6 + 2\mathcal{F}_7 + \mathcal{F}_8) \|\pi\|^2 + o(\epsilon^2) \\ &= \epsilon^2 \, \frac{3n^2 + 2n(1-4\gamma^2) - 4(1-\gamma^2)}{4n(n-1)(1-\gamma^2)} \, M_1^2 |\mathbb{S}^{n-1}| A_3 B_2 \|\pi\|^2 + o(\epsilon^2). \end{aligned} \qquad (2\text{-}18)$$

<u>Step 2</u>: energy in the half-annulus $B^N_+(0, 2\eta_0) \setminus B^N_+(0, \eta_0)$. According to (A-1), (A-3) and (A-4), see (2-6), it holds that

$$J^{\gamma}_{\hat{h}} [\psi(W_\epsilon + \Psi_\epsilon); X \setminus B^N_+(0, \eta_0)] = o(\epsilon^2). \qquad (2\text{-}19)$$

Consequently, one deduces from (2-15), (2-17)–(2-19) and Lemma B.4 that

$$J^{\gamma}_{\hat{h}} [\psi(W_\epsilon + \Psi_\epsilon); X] \le \int_{\mathbb{R}^N_+} x_N^{1-2\gamma} |\nabla W_1|^2 \, dx - \epsilon^2 \mathcal{C}(n,\gamma) |\mathbb{S}^{n-1}| A_3 B_2 \|\pi\|^2 + o(\epsilon^2) \qquad (2\text{-}20)$$

by choosing the optimal $M_1 \in \mathbb{R}$.

<u>Step 3</u>: completion of the proof. Lemma 2.2 and the fact that $\Psi_\epsilon = 0$ on M tell us that

$$\int_M |\psi(W_\epsilon + \Psi_\epsilon)|^{p+1} \, dv_{\hat{h}} = \int_{B^n(0,2\eta_0)} (\psi w_\epsilon)^{p+1} (1 + O(|\bar{x}|^3)) \, d\bar{x} \ge \int_{\mathbb{R}^n} w_1^{p+1} \, d\bar{x} + o(\epsilon^2). \qquad (2\text{-}21)$$

Combining (2-20) and (2-21) gives estimate (2-14). □

3. Umbilic conformal infinities

3A. ***Geometric background.*** For a fixed point $y \in M$ identified with $0 \in \mathbb{R}^n$, let $\bar{x} = (x_1, \dots, x_n)$ be the normal coordinate on M at y and $x_N = \rho$. The following expansion of the metric is borrowed from [Marques 2005].

Lemma 3.1. *Let $(\overline{X}, \bar{g})$ be a compact manifold with boundary $(M, \hat{h})$ and $y \in M$ such that $\pi = \pi_{;i} = \pi_{;ij} = \pi_{;ijk} = 0$, $R_{ij}[\hat{h}] = 0$ and $R_{NN}[\bar{g}] = 0$ at y. Then, in terms of Fermi coordinates around y, it holds that*

$$\begin{aligned}\sqrt{|\bar{g}|}(\bar{x}, x_N) = 1 &- \tfrac{1}{12} R_{ij;k}[\hat{h}] x_i x_j x_k - \tfrac{1}{2} R_{NN;i}[\bar{g}] x_N^2 x_i - \tfrac{1}{6} R_{NN;N}[\bar{g}] x_N^3 \\ &- \tfrac{1}{20}\big(\tfrac{1}{2} R_{ij;kl}[\hat{h}] + \tfrac{1}{9} R_{miqj}[\hat{h}] R_{mkql}[\hat{h}]\big) x_i x_j x_k x_l - \tfrac{1}{4} R_{NN;ij}[\bar{g}] x_N^2 x_i x_j \\ &- \tfrac{1}{6} R_{NN;Ni}[\bar{g}] x_N^3 x_i - \tfrac{1}{24}\big[R_{NN;NN}[\bar{g}] + 2(R_{iNjN}[\bar{g}])^2\big] x_N^4 + O(|(\bar{x}, x_N)|^5)\end{aligned} \tag{3-1}$$

and

$$\begin{aligned}\bar{g}^{ij}(\bar{x}, x_N) = \delta_{ij} &+ \tfrac{1}{3} R_{ikjl}[\hat{h}] x_k x_l + R_{iNjN}[\bar{g}] x_N^2 + \tfrac{1}{6} R_{ikjl;m}[\hat{h}] x_k x_l x_m + R_{iNjN;k}[\bar{g}] x_N^2 x_k \\ &+ \tfrac{1}{3} R_{iNjN;N}[\bar{g}] x_N^3 + \big(\tfrac{1}{20} R_{ikjl;mq}[\hat{h}] + \tfrac{1}{15} R_{iksl}[\hat{h}] R_{jmsq}[\hat{h}]\big) x_k x_l x_m x_q \\ &+ \big(\tfrac{1}{2} R_{iNjN;kl}[\bar{g}] + \tfrac{1}{3} \operatorname{Sym}_{ij}(R_{iksl}[\hat{h}] R_{sNjN}[\bar{g}])\big) x_N^2 x_k x_l + \tfrac{1}{3} R_{iNjN;kN}[\bar{g}] x_N^3 x_k \\ &+ \tfrac{1}{12}(R_{iNjN;NN}[\bar{g}] + 8 R_{iNsN}[\bar{g}] R_{sNjN}[\bar{g}]) x_N^4 + O(|(\bar{x}, x_N)|^5)\end{aligned} \tag{3-2}$$

near y (identified with a small half-ball $B_+^N(0, 2\eta_0)$ near 0 in $\mathbb{R}_+^N$). Here all tensors are computed at y and the indices m, q and s run from 1 to n as well.

To treat umbilic but nonlocally conformally flat boundaries, we also need the following extension of Lemma 2.4.

Lemma 3.2. *For $n \geq 3$, let (X^{n+1}, g^+) be an asymptotically hyperbolic manifold such that the conformal infinity $(M^n, [\hat{h}])$ is umbilic and (1-19) holds. For a fixed point $y \in M$, there exist a representative $\hat{h}$ of the class $[\hat{h}]$, the geodesic boundary-defining function ρ ($= x_N$ near M) and the metric $\bar{g} = \rho^2 g^+$ such that*

(1) $R_{ij;k}[\hat{h}](y) + R_{jk;i}[\hat{h}](y) + R_{ki;j}[\hat{h}](y) = 0$,

(2) $\operatorname{Sym}_{ijkl}\big(R_{ij;kl}[\hat{h}] + \tfrac{2}{9} R_{miqj}[\hat{h}] R_{mkql}[\hat{h}]\big)(y) = 0$,

(3) $\pi = 0$ *on* M, $\quad R_{NN;N}[\bar{g}](y) = R_{aN}[\bar{g}](y) = 0$,

(4) $R_{;ii}[\bar{g}](y) = -\dfrac{n\|W\|^2}{6(n-1)}$, $\quad R_{NN;ii}[\bar{g}](y) = -\dfrac{\|W\|^2}{12(n-1)}$, $\quad R_{iNjN}[\bar{g}](y) = R_{ij}[\bar{g}](y)$,

(5) $R_{NN;NN}[\bar{g}](y) = \dfrac{3}{2n} R_{;NN}[\bar{g}](y) - 2(R_{ij}[\bar{g}](y))^2$,

(6) $R_{iNjN;ij}[\bar{g}](y) = \dfrac{3-n}{2n} R_{;NN}[\bar{g}](y) - (R_{ij}[\bar{g}](y))^2 - \dfrac{\|W\|^2}{12(n-1)}$

if normal coordinates around $y \in (M, \hat{h})$ are assumed. Here $\|W\|$ is the norm of the Weyl tensor of $(M, \hat{h})$ at y.

Note that the first partial derivatives of $\hat{h}$ and the Christoffel symbols $\Gamma_{ij}^k[\hat{h}] = \Gamma_{ij}^k[\bar{g}]$ at y vanish. Also a simple computation utilizing $\pi = 0$ on M shows that $\Gamma_{aa}^b[\bar{g}] = \Gamma_{bN}^a[\bar{g}] = 0$ on M.

Proof of Lemma 3.2. Theorem 5.2 of [Lee and Parker 1987] guarantees the existence of a representative $\hat{h} \in [\hat{h}]$ on M such that (1), (2) and $R_{ij}[\hat{h}](y) = 0$ hold. Furthermore, [Escobar 1992b, Proposition 1.2] shows that umbilicity is preserved under the conformal transformation, and so $\pi = 0$ on M. The proof of the remaining identities in (3)–(6) is presented in two steps.

Step 1: By differentiating (2-3) in x_N and using the assumption that $\partial_N(R[g^+] + n(n+1)) = o(x_N^2)$ as $x_N \to 0$, see (1-19), we obtain

$$o(1) = n\left[\frac{\partial_N|\bar{g}|}{|\bar{g}|x_N^2} + \frac{\partial_{NN}|\bar{g}|}{|\bar{g}|x_N} - \frac{(\partial_N|\bar{g}|)^2}{|\bar{g}|^2x_N}\right] + \frac{2R[\bar{g}]}{x_N} + R_{,N}[\bar{g}] \quad \text{as } x_N \to 0. \tag{3-3}$$

Also, since we supposed that the mean curvature H vanishes on the umbilic boundary M, we get from (2-4) that $R_{NN}[\bar{g}](y) = \pi(y) = 0$. This in turn gives that $|\bar{g}|(y) = 1$ and $\partial_N|\bar{g}|(y) = \partial_{NN}|\bar{g}|(y) = R[\bar{g}](y) = 0$. Consequently, by taking the limit to y in (3-3), we find that

$$\begin{aligned} 0 &= n\left[\tfrac{1}{2}\partial_{NNN}|\bar{g}|(y) + \partial_{NNN}|\bar{g}|(y) - 0\right] + 2R_{,N}[\bar{g}](y) + R_{,N}[\bar{g}](y) \\ &= n\partial_{NNN}|\bar{g}|(y) + 2R_{,N}[\bar{g}](y). \end{aligned} \tag{3-4}$$

Now we observe from Lemma 3.1 that $\partial_{NNN}|\bar{g}|(y) = -2R_{NN;N}[\bar{g}](y)$. In addition, by the second Bianchi identity, the Codazzi equation and the fact that $\pi = 0$ on M, one can achieve

$$\begin{aligned} R_{,N}[\bar{g}] = R_{;N}[\bar{g}] &= 2R_{NN;N}[\bar{g}] + R_{ijij;N}[\bar{g}] = 2R_{NN;N}[\bar{g}] + (R_{ijiN;j}[\bar{g}] - R_{ijjN;i}[\bar{g}]) \\ &= 2R_{NN;N}[\bar{g}] + 2(\pi_{ii;jj} - \pi_{ij;ij}) = 2R_{NN;N}[\bar{g}] \end{aligned} \tag{3-5}$$

and

$$R_{iN}[\bar{g}] = \pi_{jj;i} - \pi_{ij;j} = 0$$

at $y \in M$. Combining (3-4) and (3-5), we get

$$0 = (2-n)R_{NN;N}[\bar{g}](y).$$

Since $n \geq 3$, it follows that $R_{NN;N}[\bar{g}](y) = 0$, as we wanted.

Step 2: It is well known that $R_{,ii}[\hat{h}](y) = R_{;ii}[\hat{h}](y) = -\frac{1}{6}\|W(y)\|^2$ in the normal coordinate around $y \in M$. Therefore the Gauss–Codazzi equation and the fact that $H = \pi = 0$ on M imply

$$R_{,ii}[\bar{g}](y) = 2R_{NN,ii}[\bar{g}](y) - \tfrac{1}{6}\|W(y)\|^2 \quad \text{and} \quad R_{iNjN}[\bar{g}](y) = R_{ij}[\bar{g}](y). \tag{3-6}$$

Moreover, since $\Delta_{\bar{x}}(R[g^+] + n(n+1)) = o(x_N^2)$ near $y \in \overline{X}$, refer to (1-19), by differentiating (2-3) in x_i twice, dividing the result by x_N^2 and then taking the limit to y, one obtains

$$R_{,ii}[\bar{g}](y) = 2nR_{NN,ii}[\bar{g}](y). \tag{3-7}$$

As a result, putting (3-7) into (3-6) and applying the relations at y

$$R_{;ii}[\bar{g}] = R_{,ii}[\bar{g}] \quad \text{and} \quad R_{NN;ii}[\bar{g}] = R_{NN,ii}[\bar{g}] - 2(\partial_i\Gamma_{iN}^a[\bar{g}])R_{aN}[\bar{g}] \underset{\text{by (3)}}{=} R_{NN,ii}[\bar{g}]$$

allow one to find (4).

On the other hand, arguing as before but using the hypothesis that $\partial_{NN}(R[g^+]+n(n+1)) = o(x_N^2)$ near $y \in \overline{X}$ at this time, one derives equalities

$$3R_{,NN}[\bar g](y) = -n\,\partial_{NNNN}|\bar g|(y) = 2n\big[R_{NN;NN}[\bar g](y) + 2(R_{iNjN}[\bar g](y))^2\big].$$

Because $R_{;NN}[\bar g](y) = R_{,NN}[\bar g](y)$, it is identical to (5). Hence the contracted second Bianchi identity, the Ricci identity and (3)–(5) give

$$\begin{aligned} R_{;NN}[\bar g] &= 2R_{iN;iN}[\bar g] + 2R_{NN;NN}[\bar g] = 2\big[R_{iN;Ni}[\bar g] + (R_{ij}[\bar g])^2 - (R_{aN}[\bar g])^2\big] + 2R_{NN;NN}[\bar g] \\ &= 2(R_{iN;Ni}[\bar g] + (R_{ij}[\bar g])^2) + \big(\tfrac{3}{n}R_{;NN}[\bar g] - 4(R_{ij}[\bar g])^2\big) \end{aligned}$$

at y. Now assertion (6) directly follows from the above equality and

$$\begin{aligned} R_{iN;Ni}[\bar g](y) &= R_{Njij;Ni}[\bar g](y) \\ &= -R_{iNjN;ij}[\bar g](y) + R_{NN;ii}[\bar g](y) = -R_{iNjN;ij}[\bar g](y) - \frac{\|W(y)\|^2}{12(n-1)}. \end{aligned}$$ □

3B. *Umbilic conformal infinity having the property $R_{\rho\rho;\rho}[\bar g] < 0$.* Like the previous section, we fix a smooth radial cut-off function $\psi \in C_c^\infty(\mathbb{R}^N_+)$ such that $\psi = 1$ in $B^N_+(0,\eta_0)$ and 0 in $\mathbb{R}^N_+ \setminus B^N_+(0,2\eta_0)$. Also, assume that $W_\epsilon = W_{\epsilon,0}$ denotes the bubble defined in (1-13). Let $y \in M$ be any fixed point identified with $0 \in \mathbb{R}^n$.

Lemma 3.3. *Suppose $J^\gamma_{\hat h}$ is the functional given in* (2-8). *If* (2-4) *is valid and $\pi = 0$ on M, then*

$$J^\gamma_{\hat h}[\psi W_\epsilon; B^N_+(0,\eta_0)] = \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma}|\nabla W_1|^2\,dx + \epsilon^3 R_{NN;N}[\bar g](y)\big[\tfrac14(n-2\gamma)\mathcal{F}_1' - \tfrac16\mathcal{F}_2' + \tfrac{1}{3n}\mathcal{F}_3'\big] + o(\epsilon^3) \tag{3-8}$$

for any $\epsilon > 0$ small, $n > 3+2\gamma$ and $\gamma \in (0,1)$. Here the values $\mathcal{F}_1'$, $\mathcal{F}_2'$ and $\mathcal{F}_3'$ are given in Lemma B.5.

Proof. Since $H = R_{NN}[\bar g] = 0$ at y and the bubbles W_ϵ depend only on the variables $|\bar x|$ and x_N, we have

$$\begin{aligned} \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma}|\nabla W_\epsilon|^2_{\bar g}\,dv_{\bar g} &= \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma}|\nabla W_1|^2\,dx \\ &\quad + \epsilon^3 R_{NN;N}[\bar g](y)\bigg(\frac{1}{3n}\int_{\mathbb{R}^N_+} x_N^{4-2\gamma}|\nabla_{\bar x} W_1|^2\,dx - \frac16\int_{\mathbb{R}^N_+} x_N^{4-2\gamma}|\nabla W_1|^2\,dx\bigg) + o(\epsilon^3). \end{aligned} \tag{3-9}$$

In addition, utilizing (2-5) and (3-1), we obtain

$$E(x_N) = \tfrac12(n-2\gamma)\big(R_{NN;i}[\bar g](y)x_i + \tfrac12 R_{NN;N}[\bar g](y)x_N\big)x_N^{1-2\gamma} + O(|x|^2 x_N^{1-2\gamma})$$

for $x_N \geq 0$ small enough. Therefore

$$\int_{B^N_+(0,\eta_0)} E(x_N)W_\epsilon^2\,dv_{\bar g} = \epsilon^3 R_{NN;N}[\bar g](y)\Big(\frac{n-2\gamma}{4}\Big)\int_{\mathbb{R}^N_+} x_N^{2-2\gamma}W_1^2\,dx + o(\epsilon^3). \tag{3-10}$$

Combining (3-9) and (3-10), we deduce (3-8). □

As a consequence of the previous lemma, we obtain the following result.

Proposition 3.4. *Suppose that (X^{n+1}, g^+) is an asymptotically hyperbolic manifold with umbilic conformal infinity $(M, [\hat{h}])$. If* (2-4) *is valid and $y \in M$ is a point such that $R_{NN;N}(y) < 0$, then for any $\epsilon > 0$ small and $n > 3 + 2\gamma$, we have*

$$\bar{I}^{\gamma}_{\hat{h}}[\psi W_\epsilon] \le \Lambda^{\gamma}(\mathbb{S}^n, [g_c]) + \epsilon^3 \underbrace{\frac{4n^2-12n+9-4\gamma^2}{24n(3-2\gamma)} \frac{\kappa_\gamma \int_{\mathbb{R}^N_+} x_N^{4-2\gamma} |\nabla W_1|^2 \, dx}{\left(\int_{\mathbb{R}^n} w_1^{p+1} \, dx\right)^{\frac{n-2\gamma}{n}}}}_{>0} R_{NN;N}(y) + o(\epsilon^3) < \Lambda^{\gamma}(\mathbb{S}^n, [g_c]),$$

where $\bar{I}^{\gamma}_{\hat{h}}$ is the γ-Yamabe functional given in (1-11), *and $\Lambda^{\gamma}(\mathbb{S}^n, [g_c])$ and κ_γ are the positive constants introduced in* (1-16) *and* (1-23), *respectively.*

Proof. By (A-1), (A-3) and (A-4), see (2-6), it is true that

$$J^{\gamma}_{\hat{h}}[\psi W_\epsilon; X \setminus B^N_+(0, \eta_0)] = o(\epsilon^3).$$

Moreover, we infer from (3-1) and radial symmetry of the function ψw_ϵ in $\mathbb{R}^n$ that

$$\int_M (\psi W_\epsilon)^{p+1} \, dv_{\hat{h}} = \int_{\mathbb{R}^n} w_1^{p+1} \, d\bar{x} + o(\epsilon^3).$$

Hence Lemmas 3.3 and B.5 give the desired estimate. □

3C. ***Umbilic nonlocally conformally flat conformal infinity.*** We now study the case when the boundary M is umbilic, nonlocally conformally flat and (1-19) holds. In view of Lemma 3.2(3), the tensor $R_{NN;N}[\bar{g}]$ has no role and one needs to expand the energy up to one higher order in ϵ.

Lemma 3.5. *Let $y = 0 \in M$ be any fixed point and $J^{\gamma}_{\hat{h}}$ the functional given in* (2-8). *If* (2-4) *and Lemma 3.2*(1)–(6) *are valid, then*

$$\begin{aligned} &J^{\gamma}_{\hat{h}}[\psi W_\epsilon; B^N_+(0, \eta_0)] \\ &= \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_1|^2 \, dx + \epsilon^4 \left[\frac{\|W\|^2}{4n} \left(\frac{\mathcal{F}''_5}{12(n-1)} - \frac{\mathcal{F}''_6}{2(n-1)(n+2)} - \frac{(n-2\gamma)\mathcal{F}''_4}{12(n-1)} \right) \right. \\ &\quad \left. + \frac{R_{;NN}[\bar{g}]}{2} \left(-\frac{\mathcal{F}''_2}{8n} + \frac{\mathcal{F}''_3}{4n^2} - \frac{(n-3)\mathcal{F}''_6}{n^2(n+2)} + \frac{(n-2\gamma)\mathcal{F}''_1}{4n} \right) + \frac{(R_{ij}[\bar{g}])^2}{n} \left(\frac{\mathcal{F}''_3}{2} - \frac{\mathcal{F}''_6}{n+2} \right) \right] + o(\epsilon^4) \end{aligned} \tag{3-11}$$

for any $\epsilon > 0$ small, $n > 4 + 2\gamma$ and $\gamma \in (0, 1)$. Here the tensors are computed at y and the values $\mathcal{F}''_1, \dots, \mathcal{F}''_6$ are given in Lemma B.6.

Proof. Step 1: estimate on the second- and third-order terms. To begin with, we ascertain that

$$J^{\gamma}_{\hat{h}}[\psi W_\epsilon; B^N_+(0, \eta_0)] = \int_{B^N_+(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_1|^2 \, dx + O(\epsilon^4). \tag{3-12}$$

In fact, thanks to (1-19), (2-5) and $R[\bar{g}](y) = R_{,N}[\bar{g}](y) = 0$, it holds that

$$\begin{aligned}
&\int_{B_+^N(0,\eta_0)} E(x_N) W_\epsilon^2 \, dv_{\bar{g}} \\
&\quad = \int_{B_+^N(0,\eta_0)} E(x_N) W_\epsilon^2 \, dx + O\left(\epsilon^{4+\zeta} \int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} W_1^2 |x|^{4+\zeta} \, dx\right) \\
&\quad = \epsilon^2 \Big(\frac{n-2\gamma}{4n}\Big) \int_{B_+^N(0,\eta_0/\epsilon)} x_N^{1-2\gamma} \big(R[\bar{g}](y) + \epsilon R_{,a}[\bar{g}](y) x_a + \tfrac{1}{2}\epsilon^2 R_{,ab}[\bar{g}](y) x_a x_b\big) W_1^2 \, dx + o(\epsilon^4) \\
&\quad = \epsilon^4 \Big(\frac{n-2\gamma}{4n}\Big) \cdot \Big[\frac{1}{2n} R_{;ii}[\bar{g}](y)\, \mathcal{F}_4'' + \frac{1}{2} R_{;NN}[\bar{g}](y)\, \mathcal{F}_1''\Big] + o(\epsilon^4),
\end{aligned} \tag{3-13}$$

where $\zeta > 0$ is a sufficiently small number. Because $R_{NN;N}[\bar{g}](y) = 0$ by Lemma 3.2(3), we see from (3-9) and (3-13) that estimate (3-12) is true.

<u>Step 2</u>: estimate on the fourth-order terms. Let $\sqrt{|\bar{g}|}^{(4)}$ and $(\bar{g}^{ij})^{(4)}$ be the fourth-order terms in the expansions (3-1) and (3-2) of $\sqrt{|\bar{g}|}$ and $\bar{g}^{ij}$. In view of (2-6), Lemma 3.2(2) and [Brendle 2008, Corollary 29], one can show that

$$\begin{aligned}
&\int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_\epsilon|^2 \sqrt{|\bar{g}|}^{(4)} \, dx \\
&\qquad = -\epsilon^4 \big[\tfrac{1}{4n} R_{NN;ii}[\bar{g}](y)\, \mathcal{F}_5'' + \tfrac{1}{24}(R_{NN;NN}[\bar{g}](y) + 2(R_{iNjN}[\bar{g}](y))^2)\mathcal{F}_2''\big] + o(\epsilon^4)
\end{aligned}$$

and

$$\begin{aligned}
\int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} (\bar{g}^{ij})^{(4)} \partial_i W_\epsilon \partial_j W_\epsilon \, dx = \epsilon^4 \Big[&\frac{1}{2n(n+2)}(R_{NN;ii}[\bar{g}](y) + 2R_{iNjN;ij}[\bar{g}](y))\mathcal{F}_6'' \\
&+ \frac{1}{12n}(R_{NN;NN}[\bar{g}](y) + 8(R_{iNjN}[\bar{g}](y))^2)\mathcal{F}_3''\Big] + o(\epsilon^4);
\end{aligned}$$

see [González and Wang 2017, Section 4]. Therefore (2-4), (3-9) and Lemma 3.2(4)–(6) yield

$$\begin{aligned}
&\int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_\epsilon|_{\bar{g}}^2 \, dv_{\bar{g}} \\
&= \int_{B_+^N(0,\eta_0)} x_N^{1-2\gamma} |\nabla W_1|^2 \, dx + \epsilon^4 \Big[\frac{\|W\|^2}{8n(n-1)} \Big(\frac{\mathcal{F}_5''}{6} - \frac{\mathcal{F}_6''}{n+2}\Big) + \frac{R_{;NN}[\bar{g}]}{2n} \Big(-\frac{\mathcal{F}_2''}{8} + \frac{\mathcal{F}_3''}{4n} - \frac{(n-3)\mathcal{F}_6''}{n(n+2)}\Big) \\
&\qquad + \frac{(R_{ij}[\bar{g}])^2}{n} \Big(\frac{\mathcal{F}_3''}{2} - \frac{\mathcal{F}_6''}{n+2}\Big)\Big] + o(\epsilon^4).
\end{aligned}$$

Now (3-13) and the previous estimate lead us to (3-11). □

Corollary 3.6. *Assume that (X^{n+1}, g^+) is an asymptotically hyperbolic manifold, $\hat{h}$ is the representative of the conformal infinity M in Lemma 3.1 and $\bar{I}_{\hat{h}}^\gamma$ is the γ-Yamabe functional in* (1-11). *If $n > 4 + 2\gamma$, $\gamma \in (0, 1)$ and Lemma 3.2*(1)–(6) *hold, we have*

$$\begin{aligned}
&\bar{I}_{\hat{h}}^\gamma[\psi W_\epsilon] \le \Lambda^\gamma(\mathbb{S}^n, [g_c]) + \epsilon^4 \Lambda^\gamma(\mathbb{S}^n, [g_c])^{-\frac{n-2\gamma}{2\gamma}} \kappa_\gamma\, |\mathbb{S}^{n-1}| A_3 B_2 \\
&\qquad \times \big(-\|W\|^2 \mathcal{D}_1'(n,\gamma) + R_{;NN}[\bar{g}]\mathcal{D}_2'(n,\gamma) - (R_{ij}[\bar{g}])^2 \mathcal{D}_3'(n,\gamma)\big) + o(\epsilon^4),
\end{aligned} \tag{3-14}$$

where $\Lambda^\gamma(\mathbb{S}^n,[g_c])$, κ_γ, A_3 *and* B_2 *are the positive constants introduced in* (1-16), (1-23) *and* (B-3), *respectively. Furthermore*

$$\begin{aligned}\mathcal{D}_1'(n,\gamma)&=\frac{15n^4-120n^3+20n^2(17-2\gamma^2)-80n(5-2\gamma^2)+48(4-5\gamma^2+\gamma^4)}{480n(n-1)(n-4)(n-4-2\gamma)(n-4+2\gamma)(1-\gamma^2)}>0,\\ \mathcal{D}_2'(n,\gamma)&=0 \quad \text{and} \quad \mathcal{D}_3'(n,\gamma)=\frac{5n^2-4n(13-2\gamma^2)+28(4-\gamma^2)}{5n(n-4)(n-4-2\gamma)(n-4+2\gamma)}.\end{aligned} \tag{3-15}$$

Proof. By Lemmas 3.1 and 3.2(1)–(2), it holds that

$$\begin{aligned}&\int_M (\psi W_\epsilon)^{p+1}\,dv_{\hat h}\\ &\quad=\int_{B^n(0,\eta_0)} w_\epsilon^{p+1}\bigl[1-\tfrac{1}{40}\bigl(R_{ij,kl}[\hat h]+\tfrac{2}{9}R_{miqj}[\hat h]R_{mkql}[\hat h]\bigr)x_ix_jx_kx_l+O(|\bar x|^5)\bigr]\,d\bar x+O(\epsilon^n)\\ &\quad=\int_{\mathbb{R}^n} w_1^{p+1}\,d\bar x+o(\epsilon^4).\end{aligned}$$

Thus the conclusion follows from an easy estimate,

$$J_{\hat h}^\gamma[\psi W_\epsilon; X\setminus B_+^N(0,\eta_0)]=o(\epsilon^4)$$

with Lemmas 3.5 and B.6 at once. □

It is interesting to see that the quantity $R_{;NN}[\bar g](y)$ does not contribute to the existence of a least energy solution, since the coefficient of $R_{;NN}[\bar g](y)$, denoted by $\mathcal{D}_2'(n,\gamma)$, is always zero for any n and γ. Such a phenomenon has been already observed in the boundary Yamabe problem [Marques 2005]. We also note that the number $\mathcal{D}_3'(n,\gamma)$ has a nonnegative sign in some situations: when $n=7$ and $\gamma\in\bigl[\frac12,1\bigr)$, or $n\ge 8$ and $\gamma\in(0,1)$. In order to cover lower-dimensional cases, we need a more refined test function.

Let $y\in M$ be a point such that $W[\hat h](y)\neq 0$. Motivated by [Almaraz 2010], we define functions

$$\widetilde\Psi_\epsilon=\Psi_\epsilon(\bar x,x_N)=M_2R_{iNjN}[\bar g]x_ix_jx_N^2r^{-1}\partial_rW_\epsilon=\epsilon^2\cdot\epsilon^{-\frac{n-2\gamma}{2}}\widetilde\Psi_1(\epsilon^{-1}\bar x,\epsilon^{-1}x_N)$$

for some $M_2\in\mathbb{R}$ and

$$\widetilde\Phi_\epsilon:=\psi(W_\epsilon+\widetilde\Psi_\epsilon)\quad\text{in } X.$$

Proposition 3.7. *Suppose that* (X^{n+1},g^+) *is an asymptotically hyperbolic manifold. Moreover* $\hat h$ *is the representative of the conformal infinity* M *satisfying* (2-4) *and Lemma 3.2*(1)–(6). *If* $n>4+2\gamma$ *and* $\gamma\in(0,1)$, *we have*

$$\begin{aligned}\bar I_{\hat h}^\gamma[\widetilde\Phi_\epsilon]\le\Lambda^\gamma(\mathbb{S}^n,[g_c])&+\epsilon^4\Lambda^\gamma(\mathbb{S}^n,[g_c])^{-\frac{n-2\gamma}{2\gamma}}\kappa_\gamma\,|\mathbb{S}^{n-1}|A_3B_2\\ &\times\bigl(-\|W\|^2\mathcal{D}_1(n,\gamma)+R_{;NN}[\bar g]\mathcal{D}_2(n,\gamma)-(R_{ij}[\bar g])^2\mathcal{D}_3(n,\gamma)\bigr)+o(\epsilon^4),\end{aligned}\tag{3-16}$$

where

$$\mathcal{D}_1(n,\gamma)=\mathcal{D}_1'(n,\gamma),\quad \mathcal{D}_2(n,\gamma)=0,$$

see (3-15) *for the definition of the positive constant* $\mathcal{D}_1'(n,\gamma)$, *and*

$$\mathcal{D}_3(n,\gamma)=\frac{25n^3-20n^2(9-\gamma^2)+100n(4-\gamma^2)-16(4-\gamma^2)^2}{5n(n-4-2\gamma)(n-4+2\gamma)\bigl(5n^2-4n(1+\gamma^2)-8(4-\gamma^2)\bigr)}.$$

Proof. Since $R_{NN}[\bar{g}](y) = 0$, we obtain

$$J^{\gamma}_{\hat{h}}[\widetilde{\Phi}_\epsilon; B^N_+(0,\eta_0)]$$
$$= J^{\gamma}_{\hat{h}}[\psi W_\epsilon; B^N_+(0,\eta_0)] + 2\int_{\mathbb{R}^N_+} x_N^{1-2\gamma}(\bar{g}^{ij} - \delta^{ij})\,\partial_i W_\epsilon \partial_j \widetilde{\Psi}_\epsilon\,dx + \int_{\mathbb{R}^N_+} x_N^{1-2\gamma}|\nabla\widetilde{\Psi}_\epsilon|^2\,dx + o(\epsilon^4). \quad (3\text{-}17)$$

Also a tedious computation with Lemmas 3.1 and 3.2(4) reveals that the second term of the right-hand side of (3-17) is equal to

$$\tfrac{2}{3}R_{ikjl}[\hat{h}]\int_{\mathbb{R}^N_+} x_N^{1-2\gamma}x_k x_l\,\partial_i W_\epsilon \partial_j \widetilde{\Psi}_\epsilon\,dx + 2R_{iNjN}[\bar{g}]\int_{\mathbb{R}^N_+} x_N^{3-2\gamma}\partial_i W_\epsilon \partial_j \widetilde{\Psi}_\epsilon\,dx + o(\epsilon^4)$$
$$= 0 + \epsilon^4 4M_2\left[\frac{1}{n}\mathcal{F}_3'' + \frac{1}{n(n+2)}(-\mathcal{F}_3'' + \mathcal{F}_7'')\right](R_{ij}[\bar{g}])^2 + o(\epsilon^4),$$

and it holds that

$$\int_{\mathbb{R}^N_+} x_N^{1-2\gamma}|\nabla\widetilde{\Psi}_\epsilon|^2\,dx = \epsilon^4\frac{2M_2^2}{n(n+2)}\big(\mathcal{F}_3'' - 2\mathcal{F}_7'' + \mathcal{F}_8'' + 4\mathcal{F}_6'' + 4\mathcal{F}_9'' + \mathcal{F}_{10}''\big)(R_{ij}[\bar{g}])^2 + o(\epsilon^4);$$

see (2-17) and (2-18). Here the constants $\mathcal{F}_1'', \dots, \mathcal{F}_{10}''$ are defined in Lemma B.6.

On the other hand, we have

$$J^{\gamma}_{\hat{h}}[\widetilde{\Phi}_\epsilon; X \setminus B^N_+(0,\eta_0)] = o(\epsilon^4),$$

and since $\widetilde{\Psi}_\epsilon = 0$ on M, the integral of $|\widetilde{\Phi}_\epsilon|^{p+1}$ over the boundary M does not contribute to the fourth-order term in the right-hand side of (3-16). By combining all information, employing Lemma B.6 and selecting the optimal $M_2 \in \mathbb{R}$, we complete the proof. □

One can verify that $\mathcal{D}_3(n,\gamma) > 0$ whenever $n > 4 + 2\gamma$ and $\gamma \in (0,1)$. Consequently we deduce Theorem 1.3 from Propositions 3.7 and 3.4.

4. Locally conformally flat or 2-dimensional conformal infinities

4A. ***Analysis of the Green's function.*** In this subsection, we prove Proposition 1.5. By Theorem A, solvability of problem (1-20) for each $y \in M$ is equivalent to the existence of a solution G^* to the equation

$$\begin{cases} -\operatorname{div}_{\bar{g}^*}\big((\rho^*)^{1-2\gamma}\nabla G^*(\cdot\,,y)\big) = 0 & \text{in } (X,\bar{g}^*), \\ \partial^\gamma_\nu G^*(\cdot\,,y) = \delta_y - Q^\gamma_{\hat{h}} G^*(\cdot\,,y) & \text{on } (M,\hat{h}), \end{cases}$$

and we have $|\bar{g}^*_{iN}| + |\bar{g}^*_{NN} - 1| = O(\rho^{2\gamma})$. We also recall [González and Qing 2013, Corollary 4.3] which states that if $\Lambda^\gamma(M,[\hat{h}]) > 0$, then M admits a metric $\hat{h}_0 \in [\hat{h}]$ such that $Q^\gamma_{\hat{h}_0} > 0$ on M. Thanks to the following lemma, it suffices to show Proposition 1.5 for $\hat{h}_0 \in [\hat{h}]$.

Lemma 4.1. *Let (X, g^+) be any conformally compact Einstein manifold with conformal infinity $(M,[\hat{h}])$, ρ the geodesic defining function of M in X and $\bar{g} = \rho^2 g^+$. For any positive smooth function w on M, define a new metric $\hat{h}_w = w^{4/(n-2\gamma)}\hat{h}$, denote the corresponding geodesic boundary-defining function* *by ρ_w and set $\bar{g}_w = \rho_w^2 g^+$. Suppose that $G = G(x, y)$ solves* (1-20). *Then the function*

$$G_w(x, y) := \left(\frac{\rho(x)}{\rho_w(x)}\right)^{\frac{n-2\gamma}{2}} w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, G(x, y) \quad \text{for } (x, y) \in \overline{X} \times M,\ x \neq y$$

again satisfies (1-20) *with $(\bar{g}_w, \hat{h}_w)$ and ρ_w substituted for $(\bar{g}, \hat{h})$ and ρ, respectively.*

Proof. By (1-5), the first equality in (1-20) is re-expressed as

$$L_{\bar{g}}(\rho^{\frac{1-2\gamma}{2}} G(\cdot\,, y)) + \left(\gamma^2 - \tfrac{1}{4}\right)\rho^{-(\frac{3+2\gamma}{2})} G(\cdot\,, y) = 0 \quad \text{in } (X, \bar{g}), \tag{4-1}$$

where $L_{\bar{g}}$ is the conformal Laplacian in $(X, \bar{g})$ defined in (1-3). Therefore one observes from (1-1) that G_w is a solution of (4-1) if $\bar{g}$ and ρ are replaced with $\bar{g}_w$ and ρ_w, respectively. Also, since $w = (\rho_w/\rho)^{(n-2\gamma)/2}$ on M, we see

$$\begin{aligned}
\partial_\nu^\gamma G_w(\cdot\,, y) = P^\gamma_{\hat{h}_w} G_w(\cdot\,, y) &= w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, P^\gamma_{w^{\frac{4}{n-2\gamma}}\hat{h}}((\rho/\rho_w)^{\frac{n-2\gamma}{2}}\, G(\cdot\,, y)) \\
&= w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, P^\gamma_{w^{\frac{4}{n-2\gamma}}\hat{h}}(w^{-1}\, G(\cdot\,, y)) = w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, w^{-\frac{n+2\gamma}{n-2\gamma}} P^\gamma_{\hat{h}}(G(\cdot\,, y)) \\
&= w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, w^{-\frac{n+2\gamma}{n-2\gamma}} \partial_\nu^\gamma(G(\cdot\,, y)) = w^{\frac{n+2\gamma}{n-2\gamma}}(y)\, w^{-\frac{n+2\gamma}{n-2\gamma}} \delta_y = \delta_y \quad \text{on } M,
\end{aligned}$$

where we have applied Theorem A and (1-1) for the first, fourth and fifth equalities. □

For brevity, we write $\hat{h} = \hat{h}_0$, $\bar{g} = \bar{g}^*$, $\rho = \rho^*$ and $G = G^*$ here and henceforth. Further, recalling that $Q^\gamma_{\hat{h}} > 0$ on M, let us define a norm

$$\|U\|_{\mathcal{W}^{1,q}(X,\rho^{1-2\gamma})} = \left(\int_X \rho^{1-2\gamma} |\nabla U|_{\bar{g}}^q \, dv_{\bar{g}} + \int_M Q^\gamma_{\hat{h}} U^q \, dv_{\hat{h}}\right)^{\frac{1}{q}}$$

for any $q \geq 1$ and set a space $\mathcal{W}^{1,q}(X, \rho^{1-2\gamma})$ as the completion of $C_c^\infty(\overline{X})$ with respect to the above norm.

Given any bounded Radon measure f (such as the Dirac measures), a function $U \in \mathcal{W}^{1,q}(X, \rho^{1-2\gamma})$ is said to be a *weak solution* of

$$\begin{cases} -\operatorname{div}_{\bar{g}}(\rho^{1-2\gamma}\nabla U) = 0 & \text{in } (X, \bar{g}), \\ \partial_\nu^\gamma U + Q^\gamma_{\hat{h}} U = f & \text{on } (M, \hat{h}) \end{cases} \tag{4-2}$$

if it satisfies that

$$\int_X \rho^{1-2\gamma} \langle \nabla U, \nabla \Psi \rangle_{\bar{g}} \, dv_{\bar{g}} + \int_M Q^\gamma_{\hat{h}} U \Psi \, dv_{\hat{h}} = \int_M f \Psi \tag{4-3}$$

for any $\Psi \in C^1(\overline{X})$.

The $\mathcal{W}^{1,2}(X, \rho^{1-2\gamma})$-norm is equivalent to the standard weighted Sobolev norm $\|U\|_{W^{1,2}(X,\rho^{1-2\gamma})}$; see [Choi and Kim 2017, Lemma 3.1]. Thus for any fixed $f \in (H^\gamma(M))^*$, the existence and uniqueness of a solution $U \in W^{1,2}(X, \rho^{1-2\gamma})$ to (4-2) are guaranteed by the Riesz representation theorem.

Lemma 4.2. *Assume that $n > 2\gamma$, $f \in (H^\gamma(M))^*$ and $1 \le \alpha < \min\{\frac{n}{n-2\gamma}, \frac{2n+2}{2n+1}\}$. Then there exists a constant $C = C(\overline{X}, g^+, \rho, n, \gamma, \alpha)$ such that*

$$\|U\|_{\mathcal{W}^{1,\alpha}(X,\rho^{1-2\gamma})} \le C\|f\|_{L^1(M)} \tag{4-4}$$

for a weak solution $U \in W^{1,2}(X, \rho^{1-2\gamma})$ to (4-2). *As a result, if f is the Dirac measure δ_y at $y \in M$, then* (4-2) *has a unique nonnegative weak solution $G(\cdot, y) \in \mathcal{W}^{1,\alpha}(X, \rho^{1-2\gamma})$.*

Proof. Step 1: We are going to verify estimate (4-4) by suitably modifying the argument in [Brézis and Strauss 1973, Section 5]. To this aim, we consider the *formal adjoint* of (4-2): Given any $h_0 \in L^q(M)$ and $H_1, \dots, H_N \in L^q(X, \rho^{1-2\gamma})$ for some $q > \max\{\frac{n}{2\gamma}, 2(n+1)\}$, we study a function V such that

$$\int_X \rho^{1-2\gamma}\langle\nabla V, \nabla\Psi\rangle_{\bar g}\, dv_{\bar g} + \int_M Q^\gamma_{\hat h} V\Psi\, dv_{\hat h} = \int_M h_0\Psi\, dv_{\hat h} + \sum_{a=1}^N \int_X \rho^{1-2\gamma} H_a \partial_a\Psi\, dv_{\bar g} \tag{4-5}$$

for any $\Psi \in C^1(\overline{X})$. Indeed, by the Lax–Milgram theorem, (4-5) has a unique solution $V \in W^{1,2}(X, \rho^{1-2\gamma})$. Moreover, employing Moser's iteration technique, we observe that V satisfies

$$\|V\|_{L^\infty(M)} + \|V\|_{L^\infty(X)} \le C\left(\|h_0\|_{L^q(M)} + \sum_{a=1}^N \|H_a\|_{L^q(X,\rho^{1-2\gamma})}\right). \tag{4-6}$$

Therefore taking $\Psi = V$ in (4-3) and U in (4-5) respectively (which is allowed thanks to the density argument) and then employing (4-6), we find

$$\begin{aligned}\int_M U h_0\, dv_{\hat h} + \sum_{a=1}^N \int_X \rho^{1-2\gamma}\partial_a U H_a\, dv_{\bar g} &= \int_M fV\, dv_{\hat h} \le \|f\|_{L^1(M)}\|V\|_{L^\infty(M)} \\ &\le C\|f\|_{L^1(M)}\left(\|h_0\|_{L^q(M)} + \sum_{a=1}^N \|H_a\|_{L^q(X,\rho^{1-2\gamma})}\right).\end{aligned}$$

This implies the validity of (4-4) with $\alpha = q'$, where q' designates the Hölder conjugate of q.

Step 2: Assume now that $f = \delta_y$ for some $y \in M$. Then one is capable of constructing a sequence $\{f_m\}_{m\in\mathbb{N}} \subset C^1(M)$ with an approximation to the identity or a mollifier so that $f_m \ge 0$ on M,

$$\sup_{m\in\mathbb{N}} \|f_m\|_{L^1(M)} \le C, \quad f_m \to 0 \text{ in } C^1_{\mathrm{loc}}(M \setminus \{y\}) \quad \text{and} \quad f_m \rightharpoonup \delta_y \text{ in the distributional sense.}$$

Denote by $\{U_m\}_{m\in\mathbb{N}} \subset W^{1,2}(X, \rho^{1-2\gamma})$ a sequence of the corresponding weak solutions to (4-2). By (4-4) and elliptic regularity, there exist a function $G(\cdot, y)$ and a number $\varepsilon_0 \in (0, 1)$ such that $U_m \rightharpoonup G(\cdot, y)$ weakly in $\mathcal{W}^{1,\alpha}(X, \rho^{1-2\gamma})$ and $U_m \to G(\cdot, y)$ in $C^{\varepsilon_0}_{\mathrm{loc}}(\overline{X} \setminus \{y\})$. It is a simple task to confirm that $G(\cdot, y)$ satisfies (4-3).

Also, putting $(U_m)_- \in W^{1,2}(X, \rho^{1-2\gamma})$ into (4-3) yields $U_m \ge 0$ in X, which in turn gives $G(\cdot, y) \ge 0$ in X. Finally, the uniqueness of $G(\cdot, y)$ comes as a consequence of (4-4). □

Completion of the proof of Proposition 1.5. The existence and nonnegativity of the Green's function G is deduced in the previous lemma. Owing to Hopf's lemma, see [González and Qing 2013, Theorem 3.5], G is positive on the compact manifold $\overline{X}$. Recall that the coercivity of (4-3) implies the uniqueness of G. □

4B. ***Locally conformally flat case.*** This subsection is devoted to provide the proof of Theorem 1.7, which treats locally conformally flat conformal infinities M.

Pick any point $y \in M$. Since it is supposed to be locally conformally flat, we can assume that y is the origin in $\mathbb{R}^n$ and identify a neighborhood $\mathcal{U}$ of y in M with a Euclidean ball $B^n(0, \varrho_1)$ for some $\varrho_1 > 0$ small, namely, $\hat{h}_{ij} = \delta_{ij}$ in $\mathcal{U} = B^n(0, \varrho_1)$. Write x_N to denote the geodesic defining function ρ for the boundary M near y. Then we have smooth symmetric n-tensors $h^{(1)}, \dots, h^{(n-1)}$ on $B^n(0, \varrho_1)$ such that

$$\bar{g} = h_{x_N} \oplus dx_N^2, \quad \text{where } (h_{x_N})_{ij}(\bar{x}, x_N) = \delta_{ij} + \sum_{m=1}^{n-1} h_{ij}^{(m)}(\bar{x}) x_N^m + O(x_N^n), \tag{4-7}$$

for $(\bar{x}, x_N) \in \mathcal{R}^N(\varrho_1, \varrho_2) := B^n(0, \varrho_1) \times [0, \varrho_2) \subset \overline{X}$, where $\varrho_2 > 0$ is a number small enough. In fact, as the next lemma indicates, the local conformal flatness on M and the assumption that X is Poincaré–Einstein together imply that all low-order tensors $h^{(m)}$ should vanish. In particular, the second fundamental form $h^{(1)}$ on M (up to a constant factor) is 0, which implies Remark 1.8(2).

Lemma 4.3. *If (X, g^+) is Poincaré–Einstein, we have $h^{(m)} = 0$ in* (4-7) *for each $m = 1, \dots, n-1$.*

Proof. Follow the argument of [Graham 2000], which starts from the paragraph after (2.4). Due to the condition $\hat{h}_{ij} = \delta_{ij}$, the right-hand side of (2.6) in that paper becomes 0, from which one can deduce the result. □

Therefore (4-7) is reduced to

$$\bar{g}_{ij}(\bar{x}, x_N) = \delta_{ij} + O(x_N^n) \quad \text{and} \quad |\bar{g}| = 1 + O(x_N^n) \quad \text{for } (\bar{x}, x_N) \in \mathcal{R}^N(\varrho_1, \varrho_2) \subset \overline{X}. \tag{4-8}$$

Choose any smooth function $\chi : [0, \infty) \to [0, 1]$ such that $\chi(t) = 1$ for $0 \le t \le 1$ and $\chi(t) = 0$ for $t \ge 2$. Recall the bubble W_ϵ defined in (1-13) and (1-14), the Green's function $G(\cdot, 0)$, its regular part Ψ given in (1-21), and the numbers $\alpha_{n,\gamma}$ and $g_{n,\gamma}$ given in (1-23). Then we construct a nonnegative, continuous and piecewise smooth function $\Phi_{\epsilon,\varrho_0}$ on $\overline{X}$ by

$$\Phi_{\epsilon,\varrho_0}(x) = \begin{cases} W_\epsilon(x) & \text{if } x \in X \cap B^N(0, \varrho_0), \\ V_{\epsilon,\varrho_0}(x)\big(G(x,0) - \chi_{\varrho_0}(x)\Psi(x)\big) & \text{if } x \in X \cap (B^N(0, 2\varrho_0) \setminus B^N(0, \varrho_0)), \\ V_{\epsilon,\varrho_0}(x) G(x,0) & \text{if } x \in X \setminus B^N(0, 2\varrho_0), \end{cases} \tag{4-9}$$

where $0 < \epsilon \ll \varrho_0 \le \min\{\varrho_1, \varrho_2\}/5$ sufficiently small, $\chi_{\varrho_0}(x) := \chi(|x|/\varrho_0)$ and

$$V_{\epsilon,\varrho_0}(x) := \left[\alpha_{n,\gamma}\left(\frac{\epsilon^{\frac{n-2\gamma}{2}}}{\varrho_0^{n-2\gamma}}\right) + \chi_{\varrho_0}(x)\left(W_\epsilon(x) - \alpha_{n,\gamma}\frac{\epsilon^{\frac{n-2\gamma}{2}}}{|x|^{n-2\gamma}}\right)\right] \cdot (g_{n,\gamma}\,\varrho_0^{-(n-2\gamma)} + A)^{-1}. \tag{4-10}$$

We remark that the main block V_{ϵ,ϱ_0} of the test function $\Phi_{\epsilon,\varrho_0}$ is different from the function W in (4.2) of [Escobar 1992a], but they share common characteristics such as decay properties, as proved in the next lemma.

Lemma 4.4. *There are constants $C, \eta_1, \eta_2 > 0$ depending only on n and γ such that*

$$|V_{\epsilon,\varrho_0}(x)| \le C\epsilon^{\frac{n-2\gamma}{2}} \quad \textit{for any } x \in X \setminus B^N(0, \varrho_0) \tag{4-11}$$

and

$$|\nabla_{\bar{x}} V_{\epsilon,\varrho_0}(x)| \le C\varrho_0^{-\eta_1}\epsilon^{\frac{n-2\gamma+2\eta_2}{2}} \quad \text{and} \quad |\partial_N V_{\epsilon,\varrho_0}(x)| \le C\rho_0^{-\eta_1}(\epsilon^{\frac{n-2\gamma+2\eta_2}{2}} + x_N^{2\gamma-1}\epsilon^{\frac{n+2\gamma}{2}}) \tag{4-12}$$

for $x = (\bar{x}, x_N) \in X \cap (B^N(0, 2\varrho_0) \setminus B^N(0, \varrho_0))$. *Also we have* $\nabla V_{\epsilon,\varrho_0} = 0$ *in* $X \setminus B^N(0, 2\varrho_0)$.

Proof. We observe from (A-1) and (4-10) that

$$|V_{\epsilon,\varrho_0}(x)| \le C\varrho_0^{n-2\gamma}\left[\left(\frac{\epsilon^{\frac{n-2\gamma}{2}}}{\varrho_0^{n-2\gamma}}\right) + \left|W_\epsilon(x) - \alpha_{n,\gamma}\frac{\epsilon^{\frac{n-2\gamma}{2}}}{|x|^{n-2\gamma}}\right|\right] \le C\left(\epsilon^{\frac{n-2\gamma}{2}} + \frac{\epsilon^{\frac{n-2\gamma+2\vartheta_1}{2}}}{\varrho_0^{\vartheta_1}}\right) \le C\epsilon^{\frac{n-2\gamma}{2}}$$

for all $\varrho_0 \le |x| \le 2\varrho_0$ and some $\vartheta_1 \in (0, 1)$, so (4-11) follows. One can derive (4-12) by making the use of (A-1), (A-3) and (A-4). We leave the details to the reader. □

Now we assert the following proposition, which suffices to conclude that the fractional Yamabe problem is solvable in this case.

Proposition 4.5. *For* $n > 2\gamma$ *and* $\gamma \in (0, 1)$, *let* (X^{n+1}, g^+) *be a Poincaré–Einstein manifold with conformal infinity* $(M^n, [\hat{h}])$ *such that* (1-9) *has the validity. Assume also that* M *is locally conformally flat. If Conjecture 1.6 holds and* $A > 0$, *then*

$$0 < \bar{I}_{\hat{h}}^{\gamma}[\Phi_{\epsilon,\varrho_0}] < \Lambda^\gamma(\mathbb{S}^n, [g_c]),$$

where $\bar{I}_{\hat{h}}^{\gamma}$ *is the* γ*-Yamabe functional defined in* (1-11), *and* $\Lambda^\gamma(\mathbb{S}^n, [g_c]) > 0$ *is the constant defined in* (1-16).

Proof. The proof is divided into three steps.

Step 1: estimation in $X \cap B^N(0, \varrho_0)$. Applying (1-15), (1-16), (4-8), (A-3), (A-4), Lemma A.2 and integrating by parts, we obtain

$$\begin{aligned}\kappa_\gamma &\int_{X\cap B^N(0,\varrho_0)} x_N^{1-2\gamma}|\nabla W_\epsilon|_{\bar{g}}^2\, dv_{\bar{g}} \\ &\le \Lambda^\gamma(\mathbb{S}^n, [g_c])\left(\int_{B^n(0,\varrho_0)} w_\epsilon^{p+1}\, d\bar{x}\right)^{\frac{n-2\gamma}{n}} \\ &\quad + \kappa_\gamma\int_{X\cap\partial B^N(0,\varrho_0)} x_N^{1-2\gamma}W_\epsilon\frac{\partial W_\epsilon}{\partial\nu}\, dS + \underbrace{O\left(\int_{B^n(0,\varrho_0)} x_N^{n+1-2\gamma}|\nabla W_\epsilon|^2\, d\bar{x}\right)}_{=O(\varrho_0^{2\gamma}\epsilon^{n-2\gamma})},\end{aligned} \tag{4-13}$$

where ν is the outward unit normal vector and dS is the Euclidean surface measure. On the other hand, if we write $g^+ = x_N^{-2}(dx_N^2 + h_{x_N})$, then

$$E(x_N) = -\tfrac{1}{4}(n-2\gamma)x_N^{-2\gamma}\operatorname{tr}(h_{x_N}^{-1}\partial_N h_{x_N}) = O(x_N^{n-1-2\gamma}) \tag{4-14}$$

in $X \cap B^N(0, 2\varrho_0)$; see (2-5). Therefore

$$\kappa_\gamma\int_{X\cap B^N(0,\varrho_0)} E(x_N)W_\epsilon^2\, dv_{\bar{g}} = O(\varrho_0^{2\gamma}\epsilon^{n-2\gamma}). \tag{4-15}$$

Step 2: estimation in $X \setminus B^N(0, \varrho_0)$. By its own definition (4-9) of the test function $\Phi_{\epsilon,\varrho_0}$, its energy on $\overline{X}$ can be evaluated as

$$\begin{aligned}
&\int_{X\setminus B^N(0,\varrho_0)} (\rho^{1-2\gamma}|\nabla\Phi_{\epsilon,\varrho_0}|_{\bar g}^2 + E(\rho)\Phi_{\epsilon,\varrho_0}^2)\, dv_{\bar g} \\
&\qquad = \int_{X\setminus B^N(0,\varrho_0)} \big(\rho^{1-2\gamma}\langle\nabla(V_{\epsilon,\varrho_0}^2 G), \nabla G\rangle_{\bar g} + E(\rho)V_{\epsilon,\varrho_0}^2 G^2 + \rho^{1-2\gamma}|\nabla V_{\epsilon,\varrho_0}|^2(G-\chi_{\varrho_0}\Psi)^2\big)\, dv_{\bar g} \\
&\qquad\quad + \int_{X\cap(B^N(0,2\varrho_0)\setminus B^N(0,\varrho_0))} \rho^{1-2\gamma}\big(\tfrac12\big\langle\nabla V_{\epsilon,\varrho_0}^2, \nabla(-2G\chi_{\varrho_0}\Psi + \chi_{\varrho_0}^2\Psi^2)\big\rangle_{\bar g}\big)\, dv_{\bar g} \\
&\qquad\quad + \int_{X\cap(B^N(0,2\varrho_0)\setminus B^N(0,\varrho_0))} \rho^{1-2\gamma}V_{\epsilon,\varrho_0}^2(|\nabla(\chi_{\varrho_0}\Psi)|^2 - 2\langle\nabla G, \nabla(\chi_{\varrho_0}\Psi)\rangle_{\bar g})\, dv_{\bar g} \\
&\qquad\quad + \int_{X\cap(B^N(0,2\varrho_0)\setminus B^N(0,\varrho_0))} E(\rho)V_{\epsilon,\varrho_0}^2(\chi_{\varrho_0}^2\Psi^2 - 2G\chi_{\varrho_0}\Psi)\, dv_{\bar g},
\end{aligned}$$

where $G = G(\,\cdot\,, 0)$. From (1-20), (1-22), (4-14) and Lemma 4.4, we see that

$$\begin{aligned}
&\kappa_\gamma \int_{X\setminus B^N(0,\varrho_0)} (\rho^{1-2\gamma}|\nabla\Phi_{\epsilon,\varrho_0}|_{\bar g}^2 + E(\rho)\Phi_{\epsilon,\varrho_0}^2)\, dv_{\bar g} \\
&\qquad \le -\kappa_\gamma \int_{X\cap\partial B^N(0,\varrho_0)} x_N^{1-2\gamma} V_{\epsilon,\varrho_0}^2 G\, \frac{\partial G}{\partial\nu}(1+O(x_N^n))\, dS + C\epsilon^{n-2\gamma+2\eta_2}\varrho_0^{-(n-2\gamma-2+2\eta_1)} \\
&\qquad\quad + C\epsilon^{n-2\gamma+\eta_2}\varrho_0^{\min\{1,2\gamma\}+1-\eta_1} + C\epsilon^n\varrho_0^{\min\{1,2\gamma\}+2\gamma-\eta_1} + C\epsilon^{n-2\gamma}\varrho_0^{\min\{1,2\gamma\}},
\end{aligned}\tag{4-16}$$

where $C > 0$ depends only on n, γ, ϱ_1 and ϱ_2. For instance, we have

$$\begin{aligned}
&\int_{X\setminus B^N(0,\varrho_0)} \rho^{1-2\gamma}|\nabla V_{\epsilon,\varrho_0}|^2(G-\chi_{\varrho_0}\Psi)^2\, dv_{\bar g} \\
&\qquad \le C\varrho_0^{-2\eta_1}\int_{B^N(0,2\varrho_0)\setminus B^N(0,\varrho_0)} x_N^{1-2\gamma}(\epsilon^{n-2\gamma+2\eta_2} + x_N^{2(2\gamma-1)}\epsilon^{n+2\gamma})\cdot\left(\frac{1}{|x|^{2(n-2\gamma)}}+1\right) dx \\
&\qquad \le C\big(\epsilon^{n-2\gamma+2\eta_2}\varrho_0^{-(n-2\gamma-2+2\eta_1)} + \epsilon^{n+2\gamma}\varrho_0^{-n+6\gamma}|\log\varrho_0|\big) \le C\epsilon^{n-2\gamma+2\eta_2}\varrho_0^{-(n-2\gamma-2+2\eta_1)}
\end{aligned}$$

for $0 < \epsilon \ll \varrho_0$ small. The other terms can be managed in a similar manner.

Step 3: conclusion. By combining (4-13), (4-15) and (4-16), we deduce

$$\begin{aligned}
&\kappa_\gamma \int_X (\rho^{1-2\gamma}|\nabla\Phi_{\epsilon,\varrho_0}|_{\bar g}^2 + E(\rho)\Phi_{\epsilon,\varrho_0}^2)\, dv_{\bar g} \\
&\qquad \le \Lambda^\gamma(\mathbb{S}^n, [g_c])\left(\int_{B^n(0,\varrho_0)} w_\epsilon^{p+1}\, d\bar x\right)^{\frac{n-2\gamma}{n}} + \kappa_\gamma\int_{X\cap\partial B^N(0,\varrho_0)} \underbrace{x_N^{1-2\gamma}\left(W_\epsilon\frac{\partial W_\epsilon}{\partial\nu} - V_{\epsilon,\varrho_0}^2 G\,\frac{\partial G}{\partial\nu}\right)}_{=:I} dS \\
&\qquad\quad + C\epsilon^{n-2\gamma}\varrho_0^{\min\{1,2\gamma\}}.
\end{aligned}\tag{4-17}$$

Let us compute the integral of I over the boundary $X \cap \partial B^N(0, \varrho_0)$ in the right-hand side of (4-17). Because of Lemma A.1 and (1-22), one has

$$\begin{aligned}\frac{\partial W_\epsilon}{\partial \nu} - V_{\epsilon,\varrho_0}\frac{\partial G}{\partial \nu} &\le -\frac{\alpha_{n,\gamma}(n-2\gamma)\epsilon^{\frac{n-2\gamma}{2}}}{\varrho_0^{n-2\gamma+1}} + (g_{n,\gamma}\,\varrho_0^{-(n-2\gamma)} + A)^{-1}\frac{\alpha_{n,\gamma}\,g_{n,\gamma}\,(n-2\gamma)\epsilon^{\frac{n-2\gamma}{2}}}{\varrho_0^{2(n-2\gamma)+1}} \\ &\qquad + C\epsilon^{\frac{n-2\gamma}{2}+\vartheta_1}\varrho_0^{-(n-2\gamma+1+\vartheta_1)} + C\epsilon^{\frac{n-2\gamma}{2}}\varrho_0^{\min\{0,2\gamma-1\}} \\ &\le -\alpha_{n,\gamma}\,g_{n,\gamma}^{-1}\,(n-2\gamma)A\frac{\epsilon^{\frac{n-2\gamma}{2}}}{\varrho_0} + C\epsilon^{\frac{n-2\gamma}{2}}\varrho_0^{\min\{0,2\gamma-1\}} + C\epsilon^{\frac{n-2\gamma}{2}+\vartheta_1}\varrho_0^{-(n-2\gamma+1+\vartheta_1)}\end{aligned}$$

on $\{|x| = \varrho_0\}$ for some $\vartheta_1 \in (0, 1)$. Therefore using the fact that $W_1(x) \ge \frac{1}{2}\alpha_{n,\gamma}\epsilon^{\frac{n-2\gamma}{2}}\varrho_0^{-(n-2\gamma)}$ on $\{|x| = \varrho_0\}$, we discover

$$\begin{aligned}\int_{X\cap\partial B^N(0,\varrho_0)} I\,dS &= \int_{X\cap\partial B^N(0,\varrho_0)} x_N^{1-2\gamma}\left[W_\epsilon\left(\frac{\partial W_\epsilon}{\partial \nu} - V_{\epsilon,\varrho_0}\frac{\partial G}{\partial \nu}\right) - V_{\epsilon,\varrho_0}^2\frac{\partial G}{\partial \nu}\Psi\right]dS \\ &\le -\frac{\alpha_{n,\gamma}^2}{g_{n,\gamma}}\left(\frac{n-2\gamma}{4}\right)\left(\int_{\partial B^N(0,1)} x_N^{1-2\gamma}\,dS\right)A\epsilon^{n-2\gamma} + C\epsilon^{n-2\gamma}\varrho_0^{\min\{1,2\gamma\}} \\ &\qquad + C\epsilon^{n-2\gamma+\vartheta_1}\varrho_0^{-(n-2\gamma+\vartheta_1)}.\end{aligned}$$

Now the previous estimate, (4-17), (1-16) and the assumption $A > 0$ yield that

$$\begin{aligned}\bar{I}_{\hat{h}}^\gamma[\Phi_{\epsilon,\varrho_0}] &\le \Lambda^\gamma(\mathbb{S}^n, [g_c]) - \frac{\alpha_{n,\gamma}^2}{g_{n,\gamma}}\left(\frac{S_{n,\gamma}}{\kappa_\gamma}\right)^{\frac{n-2\gamma}{2\gamma}}\left(\tfrac{1}{8}(n-2\gamma)\right)\cdot\frac{|\mathbb{S}^{n-1}|}{2}B\left(1-\gamma, \tfrac{1}{2}n\right)\cdot A\epsilon^{n-2\gamma} \\ &\qquad + C\epsilon^{n-2\gamma}\varrho_0^{\min\{1,2\gamma\}} + C\epsilon^{n-2\gamma+\vartheta_1}\varrho_0^{-(n-2\gamma+\vartheta_1)} \\ &< \Lambda^\gamma(\mathbb{S}^n, [g_c]),\end{aligned}$$

where B is the beta function. Additionally the last strict inequality holds for $0 < \epsilon \ll \varrho_0$ small enough. □

4C. *2-dimensional case.* We are now led to treat the case when $(M, [\hat{h}])$ is a 2-dimensional closed manifold.

Fix an arbitrary point $p \in M$ and let $\bar{x} = (x_1, x_2)$ be normal coordinates at p. Since X is Poincaré–Einstein, it holds that $h^{(1)} = 0$ in (4-7), whence we have

$$\bar{g}_{ij}(\bar{x}, x_N) = \delta_{ij} + O(|x|^2) \quad \text{and} \quad |\bar{g}| = 1 + O(|x|^2) \quad \text{for } (\bar{x}, x_N) \in \mathcal{R}^N(\varrho_1, \varrho_2) \subset \overline{X}, \tag{4-18}$$

where the rectangle $\mathcal{R}^N(\varrho_1, \varrho_2)$ is defined in the line following (4-7).

With Proposition B in the Introduction, the next result will give the validity of Theorem 1.7 if $n = 2$.

Proposition 4.6. *For $\gamma \in (0, 1)$, let (X^3, g^+) be a Poincaré–Einstein manifold with conformal infinity $(M^2, [\hat{h}])$ such that* (1-9) *holds. If Conjecture 1.6 holds and $A > 0$, then*

$$0 < \bar{I}_{\hat{h}}^\gamma[\Phi_{\epsilon,\varrho_0}] < \Lambda^\gamma(\mathbb{S}^2, [g_c])$$

for the test function $\Phi_{\epsilon,\varrho_0}$ introduced in (4-9).

Proof. We compute the error in $X \cap B^N_+(0, \varrho_0)$ due to the metric. As in (4-13) and (4-15), one has

$$\int_{X\cap B^N(0,\varrho_0)} x_N^{1-2\gamma} |\nabla W_\epsilon|^2_{\bar g}\, dv_{\bar g} = \int_{X\cap B^N(0,\varrho_0)} x_N^{1-2\gamma} |\nabla W_\epsilon|^2\, dx + \underbrace{O\left(\int_{X\cap B^N(0,\varrho_0)} x_N^{1-2\gamma} |x|^2 |\nabla W_\epsilon|^2\, dx\right)}_{=O(\varrho_0^{2\gamma}\epsilon^{2-2\gamma})}$$

and

$$\int_{X\cap B^N(0,\varrho_0)} E(x_N) W_\epsilon^2\, dv_{\bar g} = O\left(\int_{X\cap B^N(0,\varrho_0)} x_N^{1-2\gamma} W_\epsilon^2\, dx\right) = O(\varrho_0^{2\gamma}\epsilon^{2-2\gamma})$$

from (4-18). Therefore the error arising from the metric is ignorable, and the same argument in proof of Proposition 4.5 works. □

Appendix A: Expansion of the standard bubble $W_{1,0}$ near infinity

This appendix is devoted to finding expansions of the function $W_1 = W_{1,0}$, defined in (1-13), and its derivatives near infinity. Specifically we improve [Choi and Kim 2017, Lemma A.2] by pursuing a new approach based on conformal properties of W_1.

For the functions W_1 and $x \cdot \nabla W_1$, we have:

Lemma A.1. *Suppose that $n > 2\gamma$ and $\gamma \in (0, 1)$. For any fixed large number $R_0 > 0$, we have*

$$\left| W_1(x) - \frac{\alpha_{n,\gamma}}{|x|^{n-2\gamma}} \right| + \left| x \cdot \nabla W_1(x) + \frac{\alpha_{n,\gamma}(n-2\gamma)}{|x|^{n-2\gamma}} \right| \le \frac{C}{|x|^{n-2\gamma+\vartheta_1}} \tag{A-1}$$

for $|x| \ge R_0$, where numbers $\vartheta_1 \in (0, 1)$ and $C > 0$ rely only on n, γ and R_0.

Proof. Given any function F in $\mathbb{R}^N_+$, let F^* be its fractional Kelvin transform defined as

$$F^*(x) = \frac{1}{|x|^{n-2\gamma}}\, F\left(\frac{x}{|x|^2}\right) \quad \text{for } x \in \mathbb{R}^N_+.$$

Then it is known that $W_1^* = W_1$. Let us claim that $(x \cdot \nabla W_1)^*(0) = -\alpha_{n,\gamma}(n-2\gamma)$ and $(x \cdot \nabla W_1)^*$ is C^∞ in the $\bar x$-variable and Hölder continuous in the x_N-variable. Since

$$x_N^{2-2\gamma} \partial_{NN} W_1 = -(1-2\gamma) x_N^{1-2\gamma} \partial_N W_1 - x_N^{2-2\gamma} \Delta_{\bar x} W_1 \quad \text{in } \mathbb{R}^N_+,$$

we have

$$\begin{cases} -\operatorname{div}\big(x_N^{1-2\gamma} \nabla (x \cdot \nabla W_1)\big) = 0 & \text{in } \mathbb{R}^N_+, \\ \partial^\gamma_\nu (x \cdot \nabla W_1) = \sum_{i=1}^n x_i \partial_{x_i} \partial^\gamma_\nu W_1 + \partial^\gamma_\nu W_1 - \lim_{x_N \to 0} x_N^{2-2\gamma} \partial_{NN} W_1 & \\ \qquad = p \sum_{i=1}^n x_i \partial_{x_i}(w_1^p) + 2\gamma w_1^p & \text{on } \mathbb{R}^n. \end{cases}$$

Employing [Fall and Weth 2012, Proposition 2.6; Caffarelli and Silvestre 2007] and doing some computations, we obtain that

$$\begin{cases} -\operatorname{div}(x_N^{1-2\gamma} \nabla (x \cdot \nabla W_1)^*) = 0 & \text{in } \mathbb{R}^N_+, \\ \partial^\gamma_\nu (x \cdot \nabla W_1)^* = (-\Delta)^\gamma (x \cdot \nabla W_1)^* = \alpha^p_{n,\gamma} \left(\dfrac{2\gamma |\bar x|^2 - n}{(1 + |\bar x|^2)^{\frac{n+2\gamma+2}{2}}} \right) & \text{on } \mathbb{R}^n. \end{cases}$$

Therefore $(x\cdot\nabla W_1)^*$ has regularity stated above, and according to Green's representation formula,

$$(x\cdot\nabla W_1)^*(0)=\alpha_{n,\gamma}^p g_{n,\gamma}\int_{\mathbb{R}^n}\frac{1}{|\bar y|^{n-2\gamma}}\left(\frac{2\gamma|\bar y|^2-n}{(1+|\bar y|^2)^{\frac{n+2\gamma+2}{2}}}\right)d\bar y=-\alpha_{n,\gamma}(n-2\gamma).$$

This proves the assertion.

Now we can check (A-1) with the above observations. By standard elliptic theory, there exist constants $c_1,\dots,c_N>0$ such that

$$\left|W_1^*(x)-\alpha_{n,\gamma}\right|+\left|(x\cdot\nabla W_1)^*(x)+\alpha_{n,\gamma}(n-2\gamma)\right|\le\sum_{i=1}^n c_i|x_i|+c_N x_N^{\vartheta_1} \tag{A-2}$$

for any $|x|\le R_0^{-1}$ and some $\vartheta_1\in(0,1)$. Hence, by taking the Kelvin transform in (A-2), we see that the desired inequality (A-1) is valid for all $|x|\ge R_0$. □

Additionally we have the following decay estimate of the derivatives of W_1.

Lemma A.2. *Assume that $n>2\gamma$ and $\gamma\in(0,1)$. For any fixed large number $R_0>0$, there exist constants $C>0$ and $\vartheta_2\in(0,\min\{1,2\gamma\})$ depending only on n, γ and R_0 such that*

$$\left|\nabla_{\bar x}W_1(x)+\frac{\alpha_{n,\gamma}(n-2\gamma)\bar x}{|x|^{n-2\gamma+2}}\right|\le\frac{C}{|x|^{n-2\gamma+1+\vartheta_2}} \tag{A-3}$$

and

$$\left|\partial_N W_1(x)+\frac{\alpha_{n,\gamma}(n-2\gamma)x_N}{|x|^{n-2\gamma+2}}\right|\le C\left(\frac{1}{|x|^{n-2\gamma+2}}+\frac{x_N^{2\gamma-1}}{|x|^{n+2\gamma}}\right) \tag{A-4}$$

for $|x|\ge R_0$.

Proof. The precise values of the constants $p_{n,\gamma}$, $\alpha_{n,\gamma}$ and κ_γ, which will appear during the proof, are found in (1-23).

<u>Step 1</u>: By (1-13), (1-14) and Taylor's theorem, it holds that

$$\begin{aligned}\partial_i W_1(x)&=p_{n,\gamma}\int_{\mathbb{R}^n}\frac{1}{(|\bar y|^2+1)^{\frac{n+2\gamma}{2}}}\,\partial_i w_1(\bar x-x_N\bar y)\,d\bar y\\&=p_{n,\gamma}\int_{\mathbb{R}^n}\frac{1}{(|\bar y|^2+1)^{\frac{n+2\gamma}{2}}}\left[\partial_i w_1(-x_N\bar y)+\partial_{ij}w_1(-x_N\bar y)x_j+O(|\bar x|^2)\right]d\bar y\\&=p_{n,\gamma}\int_{\mathbb{R}^n}\frac{1}{(|\bar y|^2+1)^{\frac{n+2\gamma}{2}}}\left[\partial_{ii}w_1(0)x_i+O((x_N|\bar y|)^{\vartheta_2}|\bar x|)+O(|\bar x|^2)\right]d\bar y\\&=-\alpha_{n,\gamma}(n-2\gamma)x_i+O(|x|^{1+\vartheta_2})\end{aligned}$$

for $|x|\le R_0^{-1}$. Here we also used the facts that the $C^2(\mathbb{R}^n)$-norm of w_1 and the $C^{\vartheta_2}(\mathbb{R}^n)$-norm of $\partial_{ij}w_1$ are bounded for some $\vartheta_2\in(0,\min\{1,2\gamma\})$. On the other hand, the uniqueness of the γ-harmonic extension yields that $(\partial_i W_1)^*=\partial_i W_1$ for $i=1,\dots,n$. Therefore

$$\left|\partial_i W_1(x)+\frac{\alpha_{n,\gamma}(n-2\gamma)x_i}{|x|^{n-2\gamma+2}}\right|=\left|(\partial_i W_1)^*(x)+\alpha_{n,\gamma}(n-2\gamma)x_i^*\right|\le C(|x|^{1+\vartheta_2})^*\le\frac{C}{|x|^{n+2\gamma+1+\vartheta_2}}$$

for $|x|\ge R_0$, which is the desired inequality (A-3).

Step 2: If $\gamma = \frac{1}{2}$, it is known that

$$W_1(\bar{x}, x_N) = \alpha_{n,\frac{1}{2}} \left(\frac{1}{|\bar{x}|^2 + (x_n+1)^2} \right)^{\frac{n-1}{2}} \quad \text{for all } (\bar{x}, x_N) \in \mathbb{R}^N_+,$$

from which (A-4) follows. Therefore it is sufficient to consider when $\gamma \in (0,1) \setminus \{\frac{1}{2}\}$. In light of duality [Caffarelli and Silvestre 2007, Section 2.3], we have

$$\begin{cases} -\operatorname{div}(x_N^{1-2(1-\gamma)} \nabla (x_N^{1-2\gamma} \partial_N W_1)) = 0 & \text{in } \mathbb{R}^N_+, \\ x_N^{1-2\gamma} \partial_N W_1 = -\kappa_\gamma^{-1} w_1^p & \text{on } \mathbb{R}^n. \end{cases}$$

Hence if we define

$$F^{**}(x) = \frac{1}{|x|^{n-2(1-\gamma)}} F\left(\frac{x}{|x|^2} \right) \quad \text{for } x \in \mathbb{R}^N_+$$

for an arbitrary function F in $\mathbb{R}^N_+$, then

$$\begin{cases} -\operatorname{div}\big(x_N^{1-2(1-\gamma)} \nabla (x_N^{1-2\gamma} \partial_N W_1)^{**}\big) = 0 & \text{in } \mathbb{R}^N_+, \\ (x_N^{1-2\gamma} \partial_N W_1)^{**} = -\alpha_{n,\gamma}^p \kappa_\gamma^{-1} \dfrac{|\bar{x}|^2}{(1+|\bar{x}|^2)^{\frac{n+2\gamma}{2}}} & \text{on } \mathbb{R}^n. \end{cases}$$

This implies

$$\begin{aligned} (x_N^{1-2\gamma} \partial_N W_1)^{**}(\bar{x}, x_N) &= -\alpha_{n,\gamma}^p \kappa_\gamma^{-1} p_{n,1-\gamma}\, x_N^{2-2\gamma} \int_{\mathbb{R}^n} \frac{1}{|\bar{y}|^{n-2\gamma}} \frac{1}{(1+|\bar{y}|^2)^{\frac{n+2\gamma}{2}}}\, d\bar{y} + O(x_N^{2-2\gamma}|x| + |x|^2) \\ &= -\alpha_{n,\gamma}(n-2\gamma) x_N^{2-2\gamma} + O(x_N^{2-2\gamma}|x| + |x|^2) \end{aligned} \tag{A-5}$$

for all $|x| \le R_0^{-1}$. Accordingly, we have

$$\left| x_N^{1-2\gamma} \partial_N W_1(x) + \frac{\alpha_{n,\gamma}(n-2\gamma) x_N^{2-2\gamma}}{|x|^{n-2\gamma+2}} \right| \le C \left(\frac{x_N^{2-2\gamma}}{|x|^{n-2\gamma+3}} + \frac{1}{|x|^{n+2\gamma}} \right)$$

for $|x| \ge R_0$. Dividing the both sides by $x_N^{1-2\gamma}$ finishes the proof of (A-4). □

Appendix B: Some integrations regarding the standard bubble $W_{1,0}$ on $\mathbb{R}^N_+$

The following lemmas are due to González and Qing [2013, Section 7] and the authors [Kim et al. 2015, Section 4.3].

Lemma B.1. *Suppose that $n > 4\gamma - 1$. For each $x_N > 0$ fixed, let $\widehat{W}_1(\xi, x_N)$ be the Fourier transform of $W_1(\bar{x}, x_N)$ with respect to the variable $\bar{x} \in \mathbb{R}^n$. In addition, we use K_γ to signify the modified Bessel function of the second kind of order γ. Then we have*

$$\widehat{W}_1(\xi, x_N) = \hat{w}_1(\xi)\, \varphi(|\xi| x_N) \quad \textit{for all } \xi \in \mathbb{R}^n \textit{ and } x_N > 0,$$

where $\varphi(t) = d_1 t^\gamma K_\gamma(t)$ is the solution to

$$\phi''(t) + \frac{1-2\gamma}{t}\phi'(t) - \phi(t) = 0, \quad \phi(0) = 1 \textit{ and } \phi(\infty) = 0, \tag{B-1}$$

and $\hat{w}_1(t) := \hat{w}_1(|\xi|) = d_2|\xi|^{-\gamma} K_\gamma(|\xi|)$ *solves*

$$\phi''(t) + \frac{1+2\gamma}{t}\phi'(t) - \phi(t) = 0 \quad \text{and} \quad \lim_{t\to 0} t^{2\gamma}\phi(t) + \lim_{t\to\infty} t^{\gamma+\frac{1}{2}} e^t \phi(t) \le C \tag{B-2}$$

for some $C > 0$. *The numbers* $d_1,\ d_2 > 0$ *depend only on* n *and* γ.

Lemma B.2. *Let*

$$\begin{aligned} A_\alpha &= \int_0^\infty t^{\alpha-2\gamma}\varphi^2(t)\,dt, & B_\alpha &= \int_0^\infty t^{-\alpha+2\gamma}\hat{w}_1^2(t)t^{n-1}\,dt,\\ A'_\alpha &= \int_0^\infty t^{\alpha-2\gamma}\varphi(t)\,\varphi'(t)\,dt, & B'_\alpha &= \int_0^\infty t^{-\alpha+2\gamma}\hat{w}_1(t)\,\hat{w}'_1(t)t^{n-1}\,dt,\\ A''_\alpha &= \int_0^\infty t^{\alpha-2\gamma}(\varphi'(t))^2\,dt, & B''_\alpha &= \int_0^\infty t^{-\alpha+2\gamma}(\hat{w}'_1(t))^2 t^{n-1}\,dt \end{aligned} \tag{B-3}$$

for $\alpha \in \mathbb{N}\cup\{0\}$. *Then*

$$A_\alpha = \left(\frac{\alpha+2}{\alpha+1}\right)\cdot\left[\left(\frac{\alpha+1}{2}\right)^2 - \gamma^2\right]^{-1} A_{\alpha+2} = -\left(\frac{\alpha+1}{2}-\gamma\right)^{-1} A'_{\alpha+1} = \left(\frac{\alpha+1}{2}-\gamma\right)\left(\frac{\alpha-1}{2}+\gamma\right)^{-1} A''_\alpha$$

for α *odd*, $\alpha \ge 1$ *and*

$$\begin{aligned} B_\alpha &= \frac{4(n-\alpha+1)B_{\alpha-2}}{(n-\alpha)(n+2\gamma-\alpha)(n-2\gamma-\alpha)} = -\frac{2B'_{\alpha-1}}{n+2\gamma-\alpha},\\ B_{\alpha-2} &= \frac{(n-2\gamma-\alpha)B''_{\alpha-2}}{n+2\gamma-\alpha+2} \end{aligned}$$

for α *even*, $\alpha \ge 2$.

Proof. Apply (B-1), (B-2) and the identity

$$\int_0^\infty t^{\alpha-1}u(t)\,u'(t)\,dt = -\frac{\alpha-1}{2}\int_0^\infty t^{\alpha-2}u(t)^2\,dt,$$

which holds for any $\alpha > 1$ and $u \in C^1(\mathbb{R})$ decaying sufficiently fast. □

Utilizing the above lemmas, we compute some integrals regarding the standard bubble W_1 and its derivatives. The next identities are necessary in the energy expansion when nonminimal conformal infinities are considered. See Section 2B.

Lemma B.3. *Suppose that* $n \ge 2$ *and* $\gamma \in \left(0, \frac{1}{2}\right)$. *Then*

$$\int_{\mathbb{R}^N_+} x_N^{2-2\gamma}|\nabla W_1|^2\,dx = \frac{4}{1+2\gamma}\int_{\mathbb{R}^N_+} x_N^{2-2\gamma}(\partial_r W_1)^2\,dx = \frac{1-2\gamma}{2}\int_{\mathbb{R}^N_+} x_N^{-2\gamma}W_1^2\,dx < \infty.$$

Proof. Refer to [Choi and Kim 2017, Lemma 6.3]. □

The following is used in the energy expansion for the nonumbilic case. Refer to Sections 2C and 2D.

Lemma B.4. *For $n > 2+2\gamma$, it holds that*

$$\begin{aligned}
\mathcal{F}_1 &:= \int_{\mathbb{R}^N_+} x_N^{1-2\gamma} W_1^2\,dx = \frac{3}{2(1-\gamma^2)}|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_2 &:= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} |\nabla W_1|^2\,dx = \frac{3}{1+\gamma}|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_3 &:= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} (\partial_r W_1)^2\,dx = |\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_4 &:= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r(\partial_r W_1)(\partial_{rr} W_1)\,dx = -\tfrac{1}{2}n|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_5 &:= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2(\partial_{rr} W_1)^2\,dx = \frac{5n^3-4n(1+\gamma^2)+4(1-4\gamma^2)}{20(n-1)}|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_6 &:= \int_{\mathbb{R}^N_+} x_N^{1-2\gamma} r^2(\partial_r W_1)^2\,dx = \frac{(n+2)(3n^2-6n+4-4\gamma^2)}{8(n-1)(1-\gamma^2)}|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_7 &:= \int_{\mathbb{R}^N_+} x_N^{2-2\gamma} r^2(\partial_r W_1)(\partial_{rx_N} W_1)\,dx = -\frac{(n+2)(3n^2-6n+4-4\gamma^2)}{8(n-1)(1+\gamma)}|\mathbb{S}^{n-1}|A_3B_2,\\
\mathcal{F}_8 &:= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2(\partial_{rx_N} W_1)^2\,dx = \frac{(2-\gamma)(5n^3-4n(2-2\gamma+\gamma^2)+8(1-\gamma-2\gamma^2))}{20(n-1)(1+\gamma)}|\mathbb{S}^{n-1}|A_3B_2.
\end{aligned}$$

Here $r = |\bar{x}|$, and the positive constants A_3 and B_2 are defined by (B-3).

Proof. The values $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ and $\mathcal{F}_6$ were computed in [González and Qing 2013; Kim et al. 2015], so it suffices to consider the others.

Step 1: calculation of $\mathcal{F}_4$. Integration by parts gives

$$\begin{aligned}
\mathcal{F}_4 &= \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r(\partial_r W_1)(\partial_{rr} W_1)\,dx = |\mathbb{S}^{n-1}|\int_0^\infty x_N^{3-2\gamma}\left(\frac{1}{2}\int_0^\infty r^n\partial_r(\partial_r W_1)^2\,dr\right)dx_N\\
&= |\mathbb{S}^{n-1}|\int_0^\infty x_N^{3-2\gamma}\left(-\frac{n}{2}\int_0^\infty r^{n-1}(\partial_r W_1)^2\,dr\right)dx_N = -\frac{n}{2}\mathcal{F}_3 = -\frac{n}{2}|\mathbb{S}^{n-1}|A_3B_2.
\end{aligned}$$

Step 2: calculation of $\mathcal{F}_5$. Since $\Delta_{\bar{x}} W_1 = W_1'' + (n-1)r^{-1}W_1'$ (where $'$ stands for the differentiation in r), it holds that

$$\int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2(\Delta_{\bar{x}} W_1)^2\,dx = \mathcal{F}_5 + 2(n-1)\mathcal{F}_4 + (n-1)^2\mathcal{F}_3. \tag{B-4}$$

By the Plancherel theorem, Lemma B.1 and the relation

$$\begin{aligned}
&\Delta_\xi(|\xi|^2\hat{w}_1(|\xi|)\varphi(|\xi|x_N))\\
&= 2n\hat{w}_1\varphi + (n+2-2\gamma)|\xi|\hat{w}_1'\varphi + (n+2+2\gamma)|\xi|\hat{w}_1\varphi' x_N + |\xi|^2\hat{w}_1\varphi + 2|\xi|^2\hat{w}_1'\varphi' x_N + |\xi|^2\hat{w}_1\varphi x_N^2,
\end{aligned}$$

where the variable of $\hat{w}_1$ and $\hat{w}_1'$ is $|\xi|$, that of φ and φ' is $|\xi|x_N$, and $'$ represents the differentiation with respect to the radial variable $|\xi|$, we see

$$\begin{aligned}\int_{\mathbb{R}^N_+} & x_N^{3-2\gamma} r^2 (\Delta_{\bar{x}} W_1)^2\, dx \\ &= \int_0^\infty x_N^{3-2\gamma} \int_{\mathbb{R}^n} (-\Delta_\xi)\big(|\xi|^2 \hat{w}_1(|\xi|)\varphi(|\xi|x_N)\big) \cdot \big(|\xi|^2 \hat{w}_1(|\xi|)\varphi(|\xi|x_N)\big)\, d\xi\, dx_N \\ &= |\mathbb{S}^{n-1}|\big[2nA_3B_2 + (n+2-2\gamma)A_3B_1' + (n+2+2\gamma)A_4'B_2 + A_3B_0 + 2A_4'B_1' + A_5B_2\big].\end{aligned}$$

Therefore Lemma B.2 implies

$$\int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2 (\Delta_{\bar{x}} W_1)^2\, dx = \frac{5n^3 - 20n^2 + 4n(9-\gamma^2) - 16(1+\gamma^2)}{20(n-1)} |\mathbb{S}^{n-1}| A_3 B_2.$$

Now (B-4) and the information on $\mathcal{F}_3$ and $\mathcal{F}_4$ yield the desired estimate for $\mathcal{F}_5$.

Step 3: calculation of $\mathcal{F}_7$ and $\mathcal{F}_8$. Since the basic strategy is similar to Step 2, we will just sketch the proof. We observe

$$\begin{aligned}\mathcal{F}_7 &= \frac{1}{2}\int_0^\infty x_N^{2-2\gamma} \partial_N \left(\int_{\mathbb{R}^n} r^2 (\partial_r W_1)^2\, d\bar{x} \right) dx_N = \frac{1}{2}\int_0^\infty x_N^{2-2\gamma} \partial_N \left(\sum_{i=1}^n \int_{\mathbb{R}^n} |\bar{x}|^2 (\partial_{x_i} W_1)^2\, d\bar{x} \right) dx_N \\ &= \frac{1}{2}\int_0^\infty x_N^{2-2\gamma} \underbrace{\partial_N \left(\sum_{i=1}^n \int_{\mathbb{R}^n} (-\Delta_\xi)\big(\xi_i \hat{w}_1(|\xi|)\varphi(|\xi|x_N)\big) \cdot \big(\xi_i \hat{w}_1(|\xi|)\varphi(|\xi|x_N)\big)\, d\xi \right)}_{=(I)} dx_N.\end{aligned}$$

Owing to Lemmas B.1, B.2 and the expansion

$$\begin{aligned}(I) = -(n+1)&\int_{\mathbb{R}^n} \partial_N \big(|\xi|(\hat{w}_1\hat{w}_1')(|\xi|)\, \varphi^2(|\xi|x_N) + |\xi|\hat{w}_1^2(|\xi|)\, (\varphi\varphi')(|\xi|x_N)x_N\big)\, d\xi \\ &- \int_{\mathbb{R}^n} \partial_N \big(|\xi|^2(\hat{w}_1\hat{w}_1'')(|\xi|)\, \varphi^2(|\xi|x_N) + 2|\xi|^2(\hat{w}_1\hat{w}_1')(|\xi|)\, (\varphi\varphi')(|\xi|x_N)x_N\big)\, d\xi \\ &- \int_{\mathbb{R}^n} \partial_N \big(|\xi|^2\hat{w}_1^2(|\xi|)\, (\varphi\varphi'')(|\xi|x_N)x_N^2\big)\, d\xi,\end{aligned}$$

one can compute the integral $\mathcal{F}_7 = \frac{1}{2}\int_0^\infty x_N^{2-2\gamma}\, (I)\, dx_N$ to get its value given in the statement of the lemma. Moreover,

$$\begin{aligned}\mathcal{F}_8 &= \int_0^\infty x_N^{3-2\gamma} \left(\int_{\mathbb{R}^n} |\bar{x}|^2 |\nabla_{\bar{x}}(\partial_N W_1)|^2\, d\bar{x} \right) dx_N \\ &= \int_0^\infty x_N^{3-2\gamma} \left(\sum_{i=1}^n \int_{\mathbb{R}^n} (-\Delta_\xi)(\xi_i \partial_N \widehat{W}_1) \cdot (\xi_i \partial_N \widehat{W}_1)\, d\xi \right) dx_N.\end{aligned}$$

The rightmost term is computable with Lemmas B.1 and B.2. □

The next lemmas list the values of several integrals which are needed in the energy expansion for the umbilic case (see Sections 3B and 3C).

Lemma B.5. *For $n > 3 + 2\gamma$, let*

$$\mathcal{F}_1' := \int_{\mathbb{R}^N_+} x_N^{2-2\gamma} W_1^2 \, dx, \quad \mathcal{F}_2' := \int_{\mathbb{R}^N_+} x_N^{4-2\gamma} |\nabla W_1|^2 \, dx \quad \text{and} \quad \mathcal{F}_3' := \int_{\mathbb{R}^N_+} x_N^{4-2\gamma} (\partial_r W_1)^2 \, dx,$$

where $r = |\bar{x}|$. Then

$$\mathcal{F}_2' = \frac{3(3-2\gamma)}{2} \mathcal{F}_1' = \frac{8}{3+2\gamma} \mathcal{F}_3'.$$

Proof. One can argue as in [González and Qing 2013, Lemma 7.2] or [Choi and Kim 2017, Lemma 6.3]. □

Lemma B.6. *For $n > 4 + 2\gamma$, we have*

$$\mathcal{F}_1'' := \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} W_1^2 \, dx = \frac{4(n-3)}{(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_2'' := \int_{\mathbb{R}^N_+} x_N^{5-2\gamma} |\nabla W_1|^2 \, dx = \frac{16(n-3)(2-\gamma)}{(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_3'' := \int_{\mathbb{R}^N_+} x_N^{5-2\gamma} (\partial_r W_1)^2 \, dx = \frac{16(n-3)(4-\gamma^2)}{5(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_4'' := \int_{\mathbb{R}^N_+} x_N^{1-2\gamma} r^2 W_1^2 \, dx = \frac{n(3n^2 - 18n + 28 - 4\gamma^2)}{2(n-4)(n-4-2\gamma)(n-4+2\gamma)(1-\gamma^2)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_5'' := \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2 |\nabla W_1|^2 \, dx = \frac{n(3n^2 + 2n(-7+2\gamma) - 4(-4+3\gamma+\gamma^2))}{(n-4)(n-4-2\gamma)(n-4+2\gamma)(1+\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_6'' := \int_{\mathbb{R}^N_+} x_N^{3-2\gamma} r^2 (\partial_r W_1)^2 \, dx = \frac{(n+2)(5n^2 - 20n + 16 - 4\gamma^2)}{5(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_7'' := \int_{\mathbb{R}^N_+} x_N^{5-2\gamma} r (\partial_r W_1)(\partial_{rr} W_1) \, dx = -\frac{8n(n-3)(4-\gamma^2)}{5(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_8'' := \int_{\mathbb{R}^N_+} x_N^{5-2\gamma} r^2 (\partial_{rr} W_1)^2 \, dx = \frac{4(4-\gamma^2)(7n^3 - 14n^2 - 4n(5+\gamma^2) + 4 - 16\gamma^2)}{35(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

$$\mathcal{F}_9'' := \int_{\mathbb{R}^N_+} x_N^{4-2\gamma} r^2 (\partial_r W_1)(\partial_{r x_N} W_1) \, dx = -\frac{(n+2)(2-\gamma)(5n^2 - 20n + 16 - 4\gamma^2)}{5(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2,$$

and

$$\begin{aligned} \mathcal{F}_{10}'' &:= \int_{\mathbb{R}^N_+} x_N^{5-2\gamma} r^2 (\partial_{r x_N} W_1)^2 \, dx \\ &= \frac{4(2-\gamma)(3-\gamma)(7n^3 - 14n^2 - 4n(6 - 2\gamma + \gamma^2) + 8(2 - 3\gamma - 2\gamma^2))}{35(n-4)(n-4-2\gamma)(n-4+2\gamma)} |\mathbb{S}^{n-1}| A_3 B_2, \end{aligned}$$

where $r = |\bar{x}|$, and the positive constants A_3 and B_2 are defined by (B-3).

Proof. The proof is analogous to those of Lemma B.4 and [Kim et al. 2015, Lemma 4.4], so we skip it. □

Acknowledgements

The authors appreciate the referee for valuable comments which significantly improved the manuscript. Kim was supported by FONDECYT Grant 3140530 while he worked in Pontifical Catholic University of Chile. Musso is partially supported by FONDECYT Grant 1160135 and Millennium Nucleus Center for Analysis of PDE, NC130017. The research of Wei is partially supported by NSERC of Canada. Part of the paper was finished during the stay of Kim at the University of British Columbia and Wuhan University. He thanks the institutions for their hospitality and financial support.

Note added in the proof

(1) During the submission process, [Mayer and Ndiaye 2017] was posted on the arXiv. It proposes a proof of Theorem 1.7 without the positivity assumption on the constant A. In particular, they computed the expansion of a Green's function (compare our Conjecture 1.6 and their Corollary 6.1) and applied the Bahri–Coron-type topological argument in order to bypass the issue on A.

(2) Recently, Remarks 1.2(4) and 1.4(3) were confirmed affirmatively by the first author of this paper [Kim 2017].

(3) Suppose that $n \in \mathbb{N}$ and $\gamma \in (0, 1)$ satisfy $\mathcal{C}'(n, \gamma) > 0$, where $\mathcal{C}'(n, \gamma)$ is the quantity defined in (2-12). Moreover assume that $(M^n, [\hat{h}])$ is the conformal infinity of an asymptotic hyperbolic manifold (X, g^+) such that (1-9) and (1-18) hold, and the second fundamental form π never vanishes on M. Then the solution set of (1-2) (with $c > 0$) is compact in $C^2(M)$, as shown in [Kim et al. $\geq$ 2018].

References

[Almaraz 2010] S. d. M. Almaraz, "An existence theorem of conformal scalar-flat metrics on manifolds with boundary", *Pacific J. Math.* **248**:1 (2010), 1–22. MR Zbl

[Aubin 1976] T. Aubin, "Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire", *J. Math. Pures Appl.* (9) **55**:3 (1976), 269–296. MR Zbl

[Branson 1985] T. P. Branson, "Differential operators canonically associated to a conformal structure", *Math. Scand.* **57**:2 (1985), 293–345. MR Zbl

[Brendle 2008] S. Brendle, "Blow-up phenomena for the Yamabe equation", *J. Amer. Math. Soc.* **21**:4 (2008), 951–979. MR Zbl

[Brendle and Chen 2014] S. Brendle and S.-Y. S. Chen, "An existence theorem for the Yamabe problem on manifolds with boundary", *J. Eur. Math. Soc.* (*JEMS*) **16**:5 (2014), 991–1016. MR Zbl

[Brézis and Strauss 1973] H. Brézis and W. A. Strauss, "Semi-linear second-order elliptic equations in L^1", *J. Math. Soc. Japan* **25** (1973), 565–590. MR Zbl

[Caffarelli and Silvestre 2007] L. Caffarelli and L. Silvestre, "An extension problem related to the fractional Laplacian", *Comm. Partial Differential Equations* **32**:7-9 (2007), 1245–1260. MR Zbl

[Case 2017] J. S. Case, "Some energy inequalities involving fractional GJMS operators", *Anal. PDE* **10**:2 (2017), 253–280. MR Zbl

[Case and Chang 2016] J. S. Case and S.-Y. A. Chang, "On fractional GJMS operators", *Comm. Pure Appl. Math.* **69**:6 (2016), 1017–1061. MR Zbl

[Chang and González 2011] S.-Y. A. Chang and M. d. M. González, "Fractional Laplacian in conformal geometry", *Adv. Math.* **226**:2 (2011), 1410–1432. MR Zbl

[Chang and Yang 1995] S.-Y. A. Chang and P. C. Yang, "Extremal metrics of zeta function determinants on 4-manifolds", *Ann. of Math.* (2) **142**:1 (1995), 171–212. MR Zbl

[Chen 2009] S.-y. S. Chen, "Conformal deformation to scalar flat metrics with constant mean curvature on the boundary in higher dimensions", preprint, 2009. arXiv

[Cherrier 1984] P. Cherrier, "Problèmes de Neumann non linéaires sur les variétés riemanniennes", *J. Funct. Anal.* **57**:2 (1984), 154–206. MR Zbl

[Choi and Kim 2017] W. Choi and S. Kim, "On perturbations of the fractional Yamabe problem", *Calc. Var. Partial Differential Equations* **56**:1 (2017), art. id. 14. MR Zbl

[Djadli and Malchiodi 2008] Z. Djadli and A. Malchiodi, "Existence of conformal metrics with constant Q-curvature", *Ann. of Math.* (2) **168**:3 (2008), 813–858. MR Zbl

[Escobar 1992a] J. F. Escobar, "Conformal deformation of a Riemannian metric to a scalar flat metric with constant mean curvature on the boundary", *Ann. of Math.* (2) **136**:1 (1992), 1–50. MR Zbl

[Escobar 1992b] J. F. Escobar, "The Yamabe problem on manifolds with boundary", *J. Differential Geom.* **35**:1 (1992), 21–84. MR Zbl

[Escobar 1996] J. F. Escobar, "Conformal metrics with prescribed mean curvature on the boundary", *Calc. Var. Partial Differential Equations* **4**:6 (1996), 559–592. MR Zbl

[Fall and Weth 2012] M. M. Fall and T. Weth, "Nonexistence results for a class of fractional elliptic boundary value problems", *J. Funct. Anal.* **263**:8 (2012), 2205–2227. MR Zbl

[González and Qing 2013] M. d. M. González and J. Qing, "Fractional conformal Laplacians and fractional Yamabe problems", *Anal. PDE* **6**:7 (2013), 1535–1576. MR Zbl

[González and Wang 2017] M. d. M. González and M. Wang, "Further results on the fractional Yamabe problem: the umbilic case", *J. Geom. Anal.* (online publication March 2017).

[Graham 2000] C. R. Graham, "Volume and area renormalizations for conformally compact Einstein metrics", pp. 31–42 in *The Proceedings of the 19th Winter School "Geometry and Physics"* (Srní, 1999), edited by J. Slovák and M. Cadek, Rend. Circ. Mat. Palermo (2) Suppl. **63**, Circolo Matematico di Palermo, Palermo, 2000. MR Zbl

[Graham and Zworski 2003] C. R. Graham and M. Zworski, "Scattering matrix in conformal geometry", *Invent. Math.* **152**:1 (2003), 89–118. MR Zbl

[Guillarmou and Qing 2010] C. Guillarmou and J. Qing, "Spectral characterization of Poincaré–Einstein manifolds with infinity of positive Yamabe type", *Int. Math. Res. Not.* **2010**:9 (2010), 1720–1740. MR Zbl

[Gursky 1999] M. J. Gursky, "The principal eigenvalue of a conformally invariant differential operator, with an application to semilinear elliptic PDE", *Comm. Math. Phys.* **207**:1 (1999), 131–143. MR Zbl

[Gursky and Malchiodi 2015] M. J. Gursky and A. Malchiodi, "A strong maximum principle for the Paneitz operator and a non-local flow for the Q-curvature", *J. Eur. Math. Soc.* (*JEMS*) **17**:9 (2015), 2137–2173. MR Zbl

[Hang and Yang 2004] F. Hang and P. C. Yang, "The Sobolev inequality for Paneitz operator on three manifolds", *Calc. Var. Partial Differential Equations* **21**:1 (2004), 57–83. MR Zbl

[Hang and Yang 2015] F. Hang and P. C. Yang, "Sign of Green's function of Paneitz operators and the Q curvature", *Int. Math. Res. Not.* **2015**:19 (2015), 9775–9791. MR Zbl

[Hang and Yang 2016a] F. Hang and P. C. Yang, "Q curvature on a class of 3-manifolds", *Comm. Pure Appl. Math.* **69**:4 (2016), 734–744. MR Zbl

[Hang and Yang 2016b] F. Hang and P. C. Yang, "Q-curvature on a class of manifolds with dimension at least 5", *Comm. Pure Appl. Math.* **69**:8 (2016), 1452–1491. MR Zbl

[Humbert and Raulot 2009] E. Humbert and S. Raulot, "Positive mass theorem for the Paneitz–Branson operator", *Calc. Var. Partial Differential Equations* **36**:4 (2009), 525–531. MR Zbl

[Joshi and Sá Barreto 2000] M. S. Joshi and A. Sá Barreto, "Inverse scattering on asymptotically hyperbolic manifolds", *Acta Math.* **184**:1 (2000), 41–86. MR Zbl

[Kim 2017] S. Kim, "Conformal metrics with prescribed fractional scalar curvature on conformal infinities with positive fractional Yamabe constants", preprint, 2017. arXiv

[Kim et al. 2015] S. Kim, M. Musso, and J. Wei, “A non-compactness result on the fractional Yamabe problem in large dimensions”, preprint, 2015. arXiv

[Kim et al. $\geq$ 2018] S. Kim, M. Musso, and J. Wei, “A compactness theorem of the fractional Yamabe problem, I: The non-umbilic conformal infinity”, preprint.

[Lee and Parker 1987] J. M. Lee and T. H. Parker, “The Yamabe problem”, *Bull. Amer. Math. Soc.* (*N.S.*) **17**:1 (1987), 37–91. MR Zbl

[Marques 2005] F. C. Marques, “Existence results for the Yamabe problem on manifolds with boundary”, *Indiana Univ. Math. J.* **54**:6 (2005), 1599–1620. MR Zbl

[Marques 2007] F. C. Marques, “Conformal deformations to scalar-flat metrics with constant mean curvature on the boundary”, *Comm. Anal. Geom.* **15**:2 (2007), 381–405. MR Zbl

[Mayer and Ndiaye 2015a] M. Mayer and C. B. Ndiaye, “Barycenter technique and the Riemann mapping problem of Escobar”, preprint, 2015. arXiv

[Mayer and Ndiaye 2015b] M. Mayer and C. B. Ndiaye, “Proof of the remaining cases of the Yamabe boundary problem”, preprint, 2015. arXiv

[Mayer and Ndiaye 2017] M. Mayer and C. B. Ndiaye, “Fractional Yamabe problem on locally flat conformal infinities of Poincaré–Einstein manifolds”, preprint, 2017. arXiv

[Mazzeo and Melrose 1987] R. R. Mazzeo and R. B. Melrose, “Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature”, *J. Funct. Anal.* **75**:2 (1987), 260–310. MR Zbl

[Paneitz 2008] S. M. Paneitz, “A quartic conformally covariant differential operator for arbitrary pseudo-Riemannian manifolds (summary)”, *SIGMA Symmetry Integrability Geom. Methods Appl.* **4** (2008), art. id. 036. MR Zbl

[Qing and Raske 2006] J. Qing and D. Raske, “On positive solutions to semilinear conformally invariant equations on locally conformally flat manifolds”, *Int. Math. Res. Not.* **2006** (2006), art. id. 94172. MR Zbl

[Schoen 1984] R. Schoen, “Conformal deformation of a Riemannian metric to constant scalar curvature”, *J. Differential Geom.* **20**:2 (1984), 479–495. MR Zbl

[Schoen and Yau 1979a] R. Schoen and S. T. Yau, “On the proof of the positive mass conjecture in general relativity”, *Comm. Math. Phys.* **65**:1 (1979), 45–76. MR Zbl

[Schoen and Yau 1979b] R. Schoen and S.-T. Yau, “Proof of the positive action conjecture in quantum relativity”, *Phys. Rev. Lett.* **42**:9 (1979), 547–548.

[Schoen and Yau 1988] R. Schoen and S.-T. Yau, “Conformally flat manifolds, Kleinian groups and scalar curvature”, *Invent. Math.* **92**:1 (1988), 47–71. MR Zbl

[Trudinger 1968] N. S. Trudinger, “Remarks concerning the conformal deformation of Riemannian structures on compact manifolds”, *Ann. Scuola Norm. Sup. Pisa* (3) **22** (1968), 265–274. MR Zbl

[Yamabe 1960] H. Yamabe, “On a deformation of Riemannian structures on compact manifolds”, *Osaka Math. J.* **12**:1 (1960), 21–37. MR Zbl

Received 22 Mar 2016. Revised 10 May 2017. Accepted 10 Aug 2017.

SEUNGHYEOK KIM: shkim0401@gmail.com
Department of Mathematics, College of Natural Sciences, Hanyang University, Seoul, South Korea

MONICA MUSSO: mmusso@mat.puc.cl
Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Santiago, Chile

JUNCHENG WEI: jcwei@math.ubc.ca
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.115

ON THE FOURIER ANALYTIC STRUCTURE OF THE BROWNIAN GRAPH

JONATHAN M. FRASER AND TUOMAS SAHLSTEN

In a previous article (*Int. Math. Res. Not.* **2014**:10 (2014), 2730–2745) T. Orponen and the authors proved that the Fourier dimension of the graph of any real-valued function on $\mathbb{R}$ is bounded above by 1. This partially answered a question of Kahane (1993) by showing that the graph of the Wiener process W_t (Brownian motion) is almost surely not a Salem set. In this article we complement this result by showing that the Fourier dimension of the graph of W_t is almost surely 1. In the proof we introduce a method based on Itô calculus to estimate Fourier transforms by reformulating the question in the language of Itô drift-diffusion processes and combine it with the classical work of Kahane on Brownian images.

1. Introduction and results

1A. ***Geometric properties of Brownian motion.*** Gaussian processes are standard models in modern probability theory and perhaps the most well-studied example is the *Wiener process* (or standard *Brownian motion*) $W = W_t : \mathbb{R}^{\geqslant 0} \to \mathbb{R}$ characterised by the properties $W_0 = 0$, the map $t \mapsto W_t$ is almost surely continuous, and W_t has independent increments such that $W_t - W_s$ for $t > s$ is normally distributed:

$$W_t - W_s \sim N(0, t-s).$$

The Wiener process has far-reaching importance throughout mathematics and it is a topic of particular interest to understand its geometric structure. This can be achieved by studying several random fractals associated to the process such as *images* $W(K) := \{W_t : t \in K\}$ of compact sets $K \subset [0, \infty)$, *level sets* $L_c(W) := \{t \in \mathbb{R} : W_t = c\}$ for $c \in \mathbb{R}$, *graphs* $G(W) := \{(t, W_t) : t \in \mathbb{R}\}$ (see Figure 1) and other more delicate constructions such as SLE_κ-curves.

The basic properties of Brownian motion mean that these random fractals enjoy a certain "statistical self-similarity", which facilitates computation of their Hausdorff dimensions $\dim_{\mathrm{H}}$. Classical results include McKean's proof [1955] that $\dim_{\mathrm{H}} W(K) = \min\{1, 2\dim_{\mathrm{H}} K\}$ almost surely for each compact $K \subset [0, \infty)$. Moreover, for the level sets, $\dim_{\mathrm{H}} L_c(W) = \frac{1}{2}$ almost surely for $c = 0$ by [Taylor 1955] and for all $c \in \mathbb{R}$ by [Perkins 1981] conditioned on $L_c(W)$ being nonempty. For the Brownian graph $G(W)$, Taylor [1953] proved that $\dim_{\mathrm{H}} G(W) = \frac{3}{2}$ almost surely and Beffara [2008] computed the Hausdorff dimensions of SLE_κ-curves. Moreover, Hausdorff dimensions for similar sets given by many other Gaussian processes, such as fractional Brownian motion, have been also considered; see, for example,

Fraser acknowledges financial support from the Leverhulme Trust (RF-2016-500) and Sahlsten acknowledges the support from the European Union (ERC grant no. 306494 and Marie Skłodowska-Curie Individual Fellowship grant no. 655310).
MSC2010: primary 42B10, 60H30; secondary 11K16, 60J65, 28A80.
Keywords: Brownian motion, Wiener process, Itô calculus, Itô drift-diffusion process, Fourier transform, Fourier dimension, Salem set, graph.

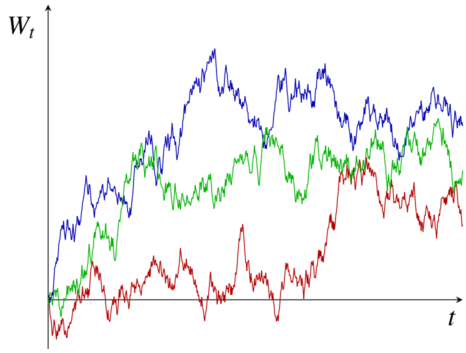

Figure 1. Three realisations of the graph $G(W)$ for the Brownian motion W_t.

Adler's classical results [1977] for fractional Brownian graphs and the recent work [Peres and Sousi 2016] concerning variable drift.

1B. ***Fourier analytic properties of Brownian motion.*** The Hausdorff dimension is the most commonly used tool for measuring the size of a set A but there is also another fundamental notion based on Fourier analysis which reveals more arithmetic and geometric features of A, including curvature, which are not seen by the Hausdorff dimension. This is based on studying the *Fourier coefficients* of a probability measure μ on $A \subset \mathbb{R}^d$, which are defined by

$$\hat{\mu}(\xi) := \int e^{-2\pi i \xi \cdot x}\, d\mu(x), \quad \xi \in \mathbb{R}^d.$$

Now the size of A can be linked to the existence of probability measures μ on A with decay of Fourier coefficients $\hat{\mu}(\xi)$ when $|\xi| \to \infty$. The following connection between Hausdorff dimension and decay of Fourier coefficients is well known and goes back to Salem and Kaufman, but we refer the reader to [Mattila 2015] for the details. If $\dim_{\mathrm{H}} A > s$, then A supports a probability measure μ with $|\hat{\mu}(\xi)| = O(|\xi|^{-s/2})$ "on average", that is, $\int_{\mathbb{R}^d} |\hat{\mu}(\xi)|^2 |\xi|^{s-d}\, d\xi < \infty$, and vice versa the Hausdorff dimension can be bounded from below if such a measure μ can be found. It is possible, however, that $\dim_{\mathrm{H}} A = s > 0$ but no measure μ on A has Fourier decay at infinity; this happens for example when A is the middle-third Cantor set in $\mathbb{R}$. Therefore, one defines the notion of *Fourier dimension* $\dim_{\mathrm{F}} A$ of a set $A \subset \mathbb{R}^d$ as the supremum of $s \in [0, d]$ for which there exists a probability measure μ supported on A such that

$$|\hat{\mu}(\xi)| = O(|\xi|^{-s/2}) \quad \text{as } |\xi| \to \infty. \tag{1-1}$$

Then by this definition we always have $\dim_{\mathrm{F}} A \leqslant \dim_{\mathrm{H}} A$ and if the two dimensions coincide then A is called a *Salem set* or a *round set* after [Kahane 1993]. In general Fourier dimension and Hausdorff dimension have no relationship other than this; in fact, Körner [2011] established that for any $0 \leqslant s < t \leqslant 1$ it is possible to construct examples $A \subset \mathbb{R}$ with $\dim_{\mathrm{F}} A = s$ and $\dim_{\mathrm{H}} A = t$. Further properties of Fourier dimension were recently developed by Ekström, Persson and Schmeling [Ekström et al. 2015]. For a more in depth account of Fourier dimension, the reader is referred to [Mattila 1995; 2015].

Finding measures μ on A with polynomially decaying Fourier transform (i.e., (1-1) for some $s > 0$) has deep links to absolute continuity, arithmetic and geometric structure, and curvature. If A supports a measure μ such that (1-1) holds with $s > 1$, then Parseval's identity yields that μ is absolutely continuous to Lebesgue measure and A must contain an interval. An application of Weyl's criterion known as the Davenport–Erdős–LeVeque criterion [Davenport et al. 1963] yields that in $\mathbb{R}$ polynomial decay of $\hat{\mu}$ guarantees that μ almost every number is normal in every base and an interesting result of Łaba and Pramanik [2009] shows that if the s in (1-1) is sufficiently close to 1 for a Frostman measure μ on $A \subset \mathbb{R}$ and there is a suitable control over the constants, see the recent work [Shmerkin 2017], then A contains nontrivial 3-term arithmetic progressions. Moreover, an analogous result also holds for higher dimensions with arithmetic patches [Chan et al. 2016].

On the curvature side, if A is a line-segment in $\mathbb{R}^2$, then A cannot contain any measure with Fourier decay at infinity so A cannot be a Salem set. However, if A is an arc of a circle or more generally a 1-dimensional smooth manifold with nonvanishing curvature then the 1-dimensional Hausdorff measure μ on A satisfies (1-1) with $s = 1$; see [Mattila 2015]. In particular, A is a Salem set. In these examples of A one can observe that the important arithmetic or curvature features present are not seen from the Hausdorff dimension.

Constructing explicit Salem sets (which are not manifolds), or just sets A supporting a measure μ satisfying (1-1) for some $s > 0$, can be achieved through, for example, Diophantine approximation by [Kaufman 1980; 1981; Bluhm 1998; Queffélec and Ramaré 2003] or via thermodynamical tools by [Jordan and Sahlsten 2016]. However, for random sets it has been observed in many instances that A is either almost surely Salem or at least supports a measure μ with (1-1) for some $s > 0$. This was first done for random Cantor sets by Salem [1951], where Salem sets were also introduced. Later in his classical papers, Kahane [1966a; 1966b] found out that the Wiener process and other Gaussian processes provide natural examples.

Since Kahane and Salem, the study of Fourier analytic properties of natural sets derived from Gaussian processes and more general random fields has been an active topic. For the Brownian images, Kahane [1985b] proved that for any compact $K \subset \mathbb{R}$ the image $W(K)$ is almost surely a Salem set of Hausdorff dimension $\min\{1, 2\dim K\}$. Kahane also established a similar result for fractional Brownian motion. Łaba and Pramanik [2009] then applied these to the additive structure of Brownian images. Later Shieh and Xiao [2006] extended Kahane's work to very general classes of Gaussian random fields. However, understanding the Fourier analytic properties of the level sets and graphs remained an important problem for some time. Kahane [1993] outlined the problem explicitly.

Problem 1.1 (Kahane). *Are the graph and level sets of a stochastic process, such as fractional Brownian motion, Salem sets*?

This precise formulation of the problem was given by Shieh and Xiao [2006, Question 2.15], but they attribute the problem to Kahane. For the Wiener process Kahane [1985a] had already established that the level sets $L_c(W)$ are Salem almost surely for any fixed $c \in \mathbb{R}$ conditioned on $L_c(W)$ being nonempty. The fractional Brownian motion case has recently been considered for $c = 0$ by Fouché and Mukeru [2013].

Kahane's problem for graphs, even in the case of the standard Brownian motion W_t, however, remained open for quite a while until, together with T. Orponen, we established that the Brownian graph $G(W)$ is almost surely *not* a Salem set [Fraser et al. 2014]. It turned out that the reason for this is purely geometric: the proof was based on the following application of a Fourier-analytic version of Marstrand's slicing lemma.

Theorem 1.2 [Fraser et al. 2014, Theorem 1.2]. *For any function $f : [0, 1] \to \mathbb{R}$ the Fourier dimension of the graph $G(f)$ cannot exceed* 1.

Indeed, since $\dim_{\mathrm{H}} G(W) = \frac{3}{2} > 1$ almost surely [Taylor 1953], this answers Kahane's problem in the negative for the Wiener process. Note that this also gives a negative answer for fractional Brownian motion since the Hausdorff dimension in that case is also strictly larger than 1 almost surely.

The methods in [Fraser et al. 2014] are purely geometric and involve no stochastic properties of Brownian motion. They also do not shed any light on the precise value for the Fourier dimension of $G(W)$. Note that even though $\dim_{\mathrm{H}} G(f) \geqslant 1$ for any continuous $f : [0, 1] \to \mathbb{R}$, the Fourier dimension of a graph may take any value in the interval [0, 1]; see [loc. cit.]. For example, $\dim_{\mathrm{F}} G(f) = 0$ if f is affine and, moreover, $\dim_{\mathrm{F}} G(f) = 0$ for the *Baire generic* $f \in C[0, 1]$; see [loc. cit., Theorem 1.3].

The main result of this paper is to complete the work initiated by Kahane's problem in the case of Brownian motion by establishing the precise almost sure value of the Fourier dimension of $G(W)$.

Theorem 1.3. *The graph $G(W)$ has Fourier dimension* 1 *almost surely.*

Moreover, the random measure μ we use to realise the Fourier dimension is Lebesgue measure dt on [0, 1] lifted onto the graph $G(W)$ via the mapping $t \mapsto (t, W_t)$. The precise estimate we obtain is that almost surely

$$|\hat{\mu}(\xi)| = O(|\xi|^{-1/2}\sqrt{\log|\xi|}) \quad \text{as } |\xi| \to \infty, \tag{1-2}$$

which combined with Theorem 1.2 yields Theorem 1.3.

A natural direction in which to continue this line of research would be to study other Gaussian processes with different covariance structure, such as the fractional Brownian motion.

1C. ***Methods: Itô calculus and reduction to Brownian images.*** The key method we introduce to estimate the Fourier transform of the graph measure μ is based on *Itô calculus*, which has previously been a natural framework in the theory of stochastic differential equations. As far as we know, Itô calculus has not been previously considered in this Fourier analytic context. Here we discuss this method and give a brief summary of the main steps in the proof. When written in polar coordinates, (1-2) asks about the rate of decay for the integral

$$\hat{\mu}(\xi) = \int_0^1 \exp(-2\pi i u(t\cos\theta + W_t \sin\theta))\,dt$$

for $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2$, $u > 0$, $\theta \in [0, 2\pi)$, as $u \to \infty$. There are two distinct cases we will consider depending on the direction of ξ, which we give a heuristic description of here.

If we ignore the random component $W_t \sin\theta$, that is, set $\theta = 0$ or π, then standard integration using the chain rule shows that $\hat{\mu}(\xi)$ equals the Fourier transform of Lebesgue measure dt at u, which decays to 0 with the polynomial rate $u^{-1} = |\xi|^{-1}$, so we are done for these directions. However, if θ is not equal to 0 or π, we still have a small random (nonsmooth) term $W_t \sin\theta$, so a classical change of variable formula or other tools from classical analysis cannot be used.

The key observation is that we can write $\hat{\mu}(\xi) = \int \exp(iX_t)\,dt$, where the stochastic process X_t satisfies the stochastic differential equation

$$dX_t := b\,dt + \sigma\,dW_t,$$

identifying it as a so-called *Itô drift-diffusion process*, where $b := -2\pi u \cos\theta$ is the *drift coefficient* of X_t and $\sigma := -2\pi u \sin\theta$ is the *diffusion coefficient* of X_t. Such processes have many useful analytic tools from Itô calculus (see Section 2) associated to them, in particular *Itô's lemma*, which works as an analogue for the chain rule. The price we pay is that Itô's lemma introduces some multiplicative error terms involving stochastic integrals, but they can be estimated with other tools from Itô calculus using moment analysis.

The estimates we obtain from Itô calculus allow us to obtain the correct Fourier decay (1-2) for μ when θ is close to 0 or π with respect to u^{-1} (more precisely, $|\sin\theta| < u^{-1/2}$), in other words, when ξ is close to pointing in the horizontal directions. Thus another estimate is needed for θ bounded away from 0 and π. This is where the classical work [Kahane 1985b] on Brownian images comes into play. If we completely ignore the deterministic component $t\cos\theta$, by setting $\theta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$, then $\hat{\mu}(\xi)$ is the Fourier transform of the Brownian image measure ν, that is, the $t \mapsto W_t$ push-forward of the Lebesgue measure dt on [0, 1] at u. Kahane [1985b] in fact already established that the decay of $|\hat{\nu}(u)|$ is almost surely of the order $u^{-1}\sqrt{\log u} = |\xi|^{-1}\sqrt{\log|\xi|}$ so (1-2) holds for these directions. A modification of Kahane's argument reveals that whenever $\theta \neq 0$ or π, then almost surely

$$|\hat{\mu}(\xi)| = O(|\sin\theta|^{-1}|\xi|^{-1}\sqrt{\log|\xi|});$$

see the discussion in Section 3C. Now one notices that when θ approaches 0 or π, this estimate blows up, and so one cannot obtain a uniform estimate over all directions from this. However, this gives (1-2) if $|\sin\theta| \geqslant u^{-1/2}$, so combining with the estimates we obtained through Itô calculus, we are done. See Section 3 for more details on the main steps of the proof.

1D. ***Other measures on the Brownian graph.*** Theorem 1.3 and (1-2) give Fourier decay for the push-forward of the Lebesgue measure on [0, 1] onto the graph $G(W)$. It would be an interesting problem to see if one can have similar results for other, possibly fractal, measures on [0, 1]. A possible problem could be:

Problem 1.4. *Classify measures τ on* [0, 1] *such that for some* $0 < s \leqslant 1$ *we have*

$$|\hat{\tau}(\xi)| = O(|\xi|^{-s/2}), \quad |\xi| \to \infty,$$

and their lift μ_τ onto the graph of $G(W)$ under $t \mapsto (t, W_t)$ satisfies

$$|\hat{\mu}_\tau(\xi)| = O(|\xi|^{-s'/2}), \quad |\xi| \to \infty$$

for any $s' < s$.

This is motivated by the fact that in [Kahane 1985b] it is possible to transfer information on the Fourier decay (or Frostman properties) of τ onto the image measure. Thus for directions θ bounded away from 0 and π we could still bound $\hat{\mu}_\tau(\xi)$ using Kahane's work. The main problem in generalising our approach to fractal measures τ on [0, 1] comes from the lack of an appropriate analogue of Itō calculus.

1E. ***Organisation of the paper.*** In Section 2 we give the necessary background from Itô calculus. In Section 3 we will give the proof of our main result Theorem 1.3. The key estimates are obtained in Section 3B and Section 3C, corresponding to the two cases discussed above.

2. Itō calculus

2A. ***Stochastic integration.*** In the proof of the main result Theorem 1.3, we end up studying integrals of the form $\int f(X_t)\,dt$ for some stochastic processes X_t and smooth scalar functions f. As standard analysis methods cannot be applied to these integrals, we need theory from stochastic analysis. Stochastic analysis provides a pleasant framework to deal with nonsmooth processes, such as the Wiener process W_t, and still preserves many of the classical features present in the smooth setting. In this section we discuss the specific tools from *Itô calculus* which we will rely on. The main references for this section are given in the book [Karatzas and Shreve 1991].

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space; that is, $\mathcal{F}_t \subset \mathcal{F}$ is an increasing filtration in t. Let $W = W_t$ be the Wiener process adapted to this filtered probability space; that is, W_t is $\mathcal{F}_t$ measurable and for each $t, s \geqslant 0$ the increment $W_{t+s} - W_t$ is independent of $\mathcal{F}_t$. We say that an $\mathbb{R}$- or $\mathbb{C}$-valued stochastic process Z_t is *adapted* if it is $\mathcal{F}_t$ measurable for all $t \geqslant 0$. We will say that a real- or complex-valued adapted process Z_t is *W_t-integrable* if the *quadratic variation* $\int_0^T |Z_t|^2\,dt$ is finite for any time $T \geqslant 0$. Given a real-valued adapted W_t-integrable stochastic process X_t, we have $\mathbb{P}$ almost surely for any time $T \geqslant 0$ it is possible to construct a *stochastic integral*

$$\int_0^T X_t\,dW_t$$

of X_t with respect to W_t in the sense of Itô; see [Karatzas and Shreve 1991, Chapter 3.2]. We use the differential notation $dU_t = X_t\,dW_t$ to mean that $\mathbb{P}$ almost surely $U_T - U_0$ is the stochastic integral of X_t with respect to W_t at time $T \geqslant 0$.

We mainly deal with complex-valued stochastic processes, so for the sake of convenience we will also introduce the *complex-valued* stochastic integral for a $\mathbb{C}$-valued W_t-integrable adapted process Z_t, defined coordinatewise using real integrals:

$$\int_0^T Z_t\,dW_t := \int_0^T \operatorname{Re} Z_t\,dW_t + i\int_0^T \operatorname{Im} Z_t\,dW_t,$$

where the real integrals are standard $\mathbb{R}$-valued stochastic integrals with respect to the Wiener process W_t. We write $dZ_t := dX_t + i\,dY_t$ for a complex-valued process $Z_t = X_t + iY_t$ with $\mathbb{R}$-valued X_t and Y_t.

2B. ***Itō drift-diffusion processes.*** The main class of adapted processes to which we apply Itô calculus is given by Wiener processes with drift and diffusion coefficients. These are called Itô drift-diffusions:

Definition 2.1 (Itô drift-diffusion process). A real- or complex-valued adapted stochastic process X_t is called an *Itô drift-diffusion process* if there exists a Lebesgue integrable adapted b_t and W_t-integrable adapted σ_t such that X_t satisfies the stochastic differential equation

$$dX_t = b_t\,dt + \sigma_t\,dW_t.$$

For Itô drift-diffusion processes there exists the following important analogue of the change of variable formula, which follows from robustness of Taylor expansions for stochastic differentials:

Lemma 2.2 (Itō's lemma). *Let X_t be an Itô drift-diffusion process and $f:\mathbb{R}\to\mathbb{R}$ twice differentiable. Then $f(X_t)$ is an Itô drift-diffusion process such that $\mathbb{P}$ almost surely for any $T\geqslant 0$ we have*

$$f(X_T)-f(X_0)=\int_0^T \bigl(b_t f'(X_t)+\tfrac{1}{2}\sigma_t^2 f''(X_t)\bigr)\,dt+\int_0^T \sigma_t f'(X_t)\,dW_t.$$

Itō's lemma was given in this pathwise form in [Karatzas and Shreve 1991, Theorem 3.3]. By using the definition of the complex-valued stochastic integral, we can also obtain a complex-valued Itô's lemma:

Lemma 2.3 (complex Itō's lemma). *Let X_t be an Itô drift-diffusion process and $f:\mathbb{R}\to\mathbb{C}$ twice differentiable. Then $f(X_t)$ is an Itô drift-diffusion process such that for $\mathbb{P}$ almost surely for any $T\geqslant 0$ we have*

$$f(X_T)-f(X_0)=\int_0^T \bigl(b_t f'(X_t)+\tfrac{1}{2}\sigma_t^2 f''(X_t)\bigr)\,dt+\int_0^T \sigma_t f'(X_t)\,dW_t.$$

Proof. We can write $f=f_1+if_2$ for real-valued twice differentiable $f_1,\,f_2:\mathbb{R}\to\mathbb{R}$. Then the derivatives satisfy $f'=f_1'+if_2'$ and $f''=f_1''+if_2''$. Moreover, by Itô's lemma (Lemma 2.2) we obtain for each $j=1,2$ that

$$df_j(X_t)=\bigl(b_t f_j'(X_t)+\tfrac{1}{2}\sigma_t^2 f_j'(X_t)\bigr)\,dt+\sigma_t f_j'(X_t)\,dW_t.$$

Then by the convention $df(X_t)=df_1(X_t)+i\,df_2(X_t)$ this gives

$$df(X_t)=\bigl(b_t f'(X_t)+\tfrac{1}{2}\sigma_t^2 f''(X_t)\bigr)\,dt+\sigma_t f'(X_t)\,dW_t$$

as required. □

2C. *Moment estimation.* Itô's lemma allows us to pass from integrals of the form $\int_0^T f(X_t)\,dt$ to $\int_0^T g(X_t)\,dW_t$ for functions g obtained from derivatives of f. In our case we will end up trying to understand the higher moments of the stochastic integrals $\int_0^T g(X_t)\,dW_t$, which will tell us about the distribution of these integrals. A very standard tool to compute the moments in Itô calculus are the *Itô isometry* and more general *Burkholder–Davis–Gundy inequalities* [Burkholder et al. 1972], which allow us to pass from stochastic integrals to their quadratic variations (that just involve Lebesgue integral).

Lemma 2.4 (Burkholder–Davis–Gundy inequality). *Let X_t be a real-valued W_t-integrable adapted process. Then for all $1\leqslant p<\infty$ we have*

$$\mathbb{E}\biggl[\biggl(\sup_{0\leqslant s\leqslant 1}\biggl|\int_0^s X_t\,dW_t\biggr|\biggr)^{2p}\biggr]\leqslant 2\sqrt{10p}\;\mathbb{E}\biggl[\biggl(\int_0^1 X_t^2\,dt\biggr)^{p}\biggr].$$

This version with the constant $2\sqrt{10p}$ was given by Peškir [1996].

3. Proof of the main result

3A. ***Preliminaries and overview of the proof.*** Let us now review how we will prove (1-2) and thus Theorem 1.3. Fix $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2$ with modulus $u > 0$ and argument $\theta \in [0, 2\pi)$. Notice that by the definition of the graph measure μ, the Fourier transform has the form

$$\hat{\mu}(\xi) = \int_0^1 \exp(iX_t)\, dt,$$

where X_t is the real-valued stochastic process

$$X_t := -2\pi u(t\cos\theta + W_t \sin\theta). \tag{3-1}$$

The first observation is that X_t is an adapted W_t-integrable process and in fact an Itô drift-diffusion process (recall Definition 2.1) satisfying

$$dX_t = b\,dt + \sigma\, dW_t$$

for deterministic and time independent coefficients $b = -2\pi u\cos\theta$ and $\sigma = -2\pi u \sin\theta$. The proof of bounding $\hat{\mu}(\xi)$ will heavily depend on the value of the angle θ we have for ξ and in particular how close the determining angle θ is to 0, π or 2π with respect to $u^{-1/2}$. For this purpose, we define the notions of *horizontal* and *vertical* angles:

Definition 3.1 (horizontal and vertical angles). Define the threshold angle

$$\theta_u := \min\{u^{-1/2}, \tfrac{\pi}{4}\}.$$

Partition the angles $[0, 2\pi)$ using θ_u into the *horizontal angles*

$$H_u := [0, \theta_u] \cup [\pi - \theta_u, \pi + \theta_u] \cup [2\pi - \theta_u, 2\pi)$$

and the *vertical angles*

$$V_u := [0, 2\pi) \setminus H_u.$$

In other words H_u contains the θ_u neighborhoods of 0 and π on the circle mod 2π and V_u the $\frac{\pi}{2} - \theta_u$ neighborhoods of $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ respectively; see Figure 2.

The proof will split into two cases in Sections 3B and 3C for bounding the Fourier transform $\hat{\mu}(\xi)$ depending on whether $\theta \in H_u$ or $\theta \in V_u$:

(1) Section 3B concerns angles $\theta \in H_u$, that is, close to horizontal directions 0 or π, and as mentioned in the Introduction our main hope here is that the smallness (with respect to $u^{-1/2}$) of the diffusion component bW_t will help us in transferring the decay of Lebesgue measure to the decay of $\hat{\mu}$. This is where Itô's lemma (see Lemma 2.3) becomes crucial as it can be applied to the process $f(X_t)$ with the function $f(x) = \exp(ix)$.

(2) Section 3C handles the angles $\theta \in V_u$ and here the plan is to use the fact that we are $u^{-1/2}$-bounded away from horizontal angles to ignore the drift component bt of the drift-diffusion process X_t and apply Kahane's bound for these directions. This turns out to be possible due to a representation of the higher moments Kahane obtained in his result on Brownian images.

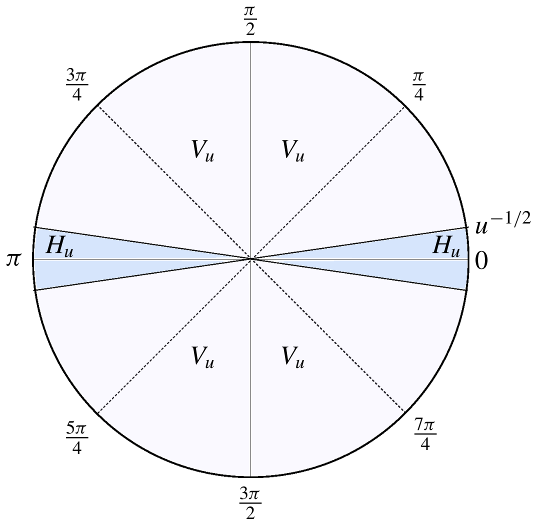

Figure 2. Splitting of $[0, 2\pi)$ to horizontal angles H_u and the vertical angles V_u.

It turns out that in both Sections 3B and 3C we only obtain decay of the Fourier transform $\hat{\mu}(\boldsymbol{k})$ for $\boldsymbol{k}$ in an ε-grid $\varepsilon\mathbb{Z}^2$ for all small $\varepsilon > 0$. Here the randomness will depend on $\varepsilon > 0$ but thanks to an argument also used by Kahane [1985b], one can pass from this information to the full decay almost surely. See Section 3D for the details.

Let us now proceed to bound $|\hat{\mu}(\xi)|$. In both Sections 3B and 3C below we will end up bounding trigonometric functions with respect to θ_u and for this purpose we will need the following standard bounds, which we record here for convenience:

Lemma 3.2 (trigonometric bounds). *We have the following bounds*:

(1) *If* $\theta \in H_u$, *then*

$$|\sin\theta| \leqslant u^{-1/2} \quad \textit{and} \quad |\cos\theta| \geqslant \tfrac{1}{\sqrt{2}}.$$

(2) *If* $\theta \in V_u$, *then*

$$|\sin\theta| \geqslant \min\left\{\tfrac{2}{\pi}u^{-1/2}, \tfrac{1}{\sqrt{2}}\right\}.$$

Proof. For $\alpha \in \left[0, \frac{\pi}{2}\right]$ we have that both cosine and sine are nonnegative. Moreover, here $\frac{2}{\pi}\alpha \leqslant \sin\alpha \leqslant \alpha$. Thus for $\theta \in [0, \theta_u]$ we have

$$\sin\theta \leqslant \theta \leqslant \theta_u \leqslant u^{-1/2} \quad \text{and} \quad \cos\theta \geqslant \cos\theta_u \geqslant \cos\tfrac{\pi}{4} = \tfrac{1}{\sqrt{2}}.$$

and for $\theta \in \left(\theta_u, \frac{\pi}{2}\right]$ as $\sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}$ we obtain

$$\sin\theta \geqslant \min\left\{\tfrac{2}{\pi}u^{-1/2}, \tfrac{1}{\sqrt{2}}\right\}.$$

This gives the claim as we may reduce the estimates back to the estimates for $\theta \in \left[0, \frac{\pi}{2}\right]$ by using standard invariance identities for sine and cosine. $\square$

3B. *Horizontal angles.* When $\theta \in H_u$ we will first obtain the following estimate on ε-grids:

Lemma 3.3. *Fix $\varepsilon > 0$. Almost surely there exists a random constant $C_\omega > 0$ such that for any $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in H_u$ we have*

$$|\hat{\mu}(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2}.$$

Given $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\}$ and a realisation (W_t), define a random variable $T = T_\omega(\xi) \in [0, 1]$, to be the minimum value of $t \in [0, 1]$ such that

$$X_t = \begin{cases} -2\pi \lceil u(\cos\theta + W_1 \sin\theta) \rceil & \text{if } X_1 \geqslant 0, \\ -2\pi \lfloor u(\cos\theta + W_1 \sin\theta) \rfloor & \text{if } X_1 < 0. \end{cases}$$

Such a time T exists almost surely since $X_0 = 0$ and X_t is almost surely continuous (since W_t is almost surely continuous). Splitting the integral of Z_t up into "complete rotations" and "what is left over", one obtains

$$\int_0^1 Z_t \, dt = \int_0^T Z_t \, dt + \int_T^1 Z_t \, dt.$$

For the integral over $[T, 1]$ we get the following estimate.

Lemma 3.4. *Almost surely there exists a random constant $C_\omega > 0$ such that for any $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\}$ with $\theta \in H_u$ we have*

$$\left| \int_T^1 Z_t \, dt \right| \leqslant C_\omega |\xi|^{-1/2}.$$

Proof. Since W_t is almost surely continuous, there almost surely exists a random constant $M_\omega > 1$ such that $W_t \in [-M_\omega, M_\omega]$ for all $t \in [0, 1]$. Define the real-valued process

$$Y_t := u(t\cos\theta + W_t \sin\theta),$$

so $X_t = -2\pi Y_t$. Suppose $X_1 \geqslant 0$. In this case $Y_T = \lceil Y_1 \rceil \leqslant 0$ and so $Y_1 + 1 \geqslant Y_T \geqslant Y_1$. Moreover, when $X_1 < 0$ we have $Y_T = \lfloor Y_1 \rfloor > 0$ and $Y_1 \geqslant Y_T \geqslant Y_1 - 1$. Thus no matter what the sign of X_1 is, we always have almost surely

$$u(\cos\theta + W_1 \sin\theta) + 1 \geqslant u(T\cos\theta + W_T \sin\theta) \geqslant u(\cos\theta + W_1 \sin\theta) - 1.$$

Therefore, in the case $\cos\theta > 0$ we obtain

$$T \geqslant 1 + W_1 \frac{\sin\theta}{\cos\theta} - W_T \frac{\sin\theta}{\cos\theta} - \frac{1}{u\cos\theta},$$

and when $\cos\theta < 0$ we have

$$T \geqslant 1 + W_1 \frac{\sin\theta}{\cos\theta} - W_T \frac{\sin\theta}{\cos\theta} + \frac{1}{u\cos\theta}.$$

Since $u \in H_u$, Lemma 3.2 together with $W_t \in [-M_\omega, M_\omega]$ yields

$$T \geqslant 1 - 2\sqrt{2} M_\omega u^{-1/2} - \frac{\sqrt{2}}{u}.$$

Recalling $M_\omega > 1$ this gives

$$\left|\int_T^1 Z_t\,dt\right| \leqslant \int_T^1 |Z_t|\,dt = 1 - T \leqslant 4M_\omega u^{-1/2}$$

as required. $\square$

We now estimate the integral over $[0, T]$, which is where Itô calculus comes into play.

Lemma 3.5. *Fix $\varepsilon > 0$. Almost surely there exists a random constant $C_\omega > 0$ such that for any $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in H_u$ we have*

$$\left|\int_0^T Z_t\,dt\right| \leqslant C_\omega |\boldsymbol{k}|^{-1/2}.$$

To prove Lemma 3.5, we first need to compute the higher-order moments of the random variable $\int_0^T Z_t\,dt$.

Lemma 3.6. *For any $p \in \mathbb{N}$ and $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\}$ with $\theta \in H_u$, the $(2p)$-th moment satisfies*

$$\mathbb{E}\left|\int_0^T Z_t\,dt\right|^{2p} \leqslant 13p^{1/2}\,4^p|\xi|^{-p}.$$

Proof. Recall that

$$X_t = -2\pi u(t\cos\theta + W_t \sin\theta)$$

is an Itô drift-diffusion process satisfying the stochastic differential equation

$$dX_t = b\,dt + \sigma\,dW_t$$

for deterministic and time-independent coefficients $b = -2\pi u\cos\theta$ and $\sigma = -2\pi u \sin\theta$. Writing

$$f(x) := \exp(ix), \quad x \in \mathbb{R},$$

we have $Z_t = f(X_t)$, $f'(x) = i\exp(ix)$ and $f''(x) = -\exp(ix)$. Thus by complex Itô's lemma (see Lemma 2.3) we have $\mathbb{P}$ almost surely

$$f(X_T) - f(X_0) = (bi - \sigma^2/2)\int_0^T f(X_t)\,dt + \sigma i \int_0^T f(X_t)\,dW_t. \tag{3-2}$$

Note that $T_\omega \leqslant 1$ is random and only $\mathcal{F}_1$ measurable; thus it is not a stopping time. However, as Lemma 2.3 is given *pathwise*, that is, $\mathbb{P}$ almost surely Itô's lemma holds for any time $T \geqslant 0$, then as T_ω is $\mathbb{P}$ almost surely well-defined, we have (3-2) almost surely. Since X_0 and X_T are 2π multiples of integers by definition, we have $f(X_T) = f(X_0) = 1$. Thus (3-2) gives

$$\int_0^T f(X_t)\,dt = -\frac{\sigma i}{bi - \sigma^2/2}\int_0^T f(X_t)\,dW_t$$

Since b and σ are deterministic, this yields that the $(2p)$-th moment satisfies

$$\mathbb{E}\left|\int_0^T f(X_t)\,dt\right|^{2p} = \left|\frac{\sigma i}{bi - \sigma^2/2}\right|^{2p}\mathbb{E}\left|\int_0^T f(X_t)\,dW_t\right|^{2p}.$$

Applying the Burkholder–Davis–Gundy inequality (see Lemma 2.4) for the process $\cos X_t$ gives

$$\mathbb{E}\left[\left|\int_0^T \cos X_t\, dW_t\right|^{2p}\right] \leqslant \mathbb{E}\left[\left(\sup_{0\leqslant s\leqslant 1}\left|\int_0^s \cos X_t\, dW_t\right|\right)^{2p}\right] \leqslant 2\sqrt{10p}\, \mathbb{E}\left(\int_0^1 \cos^2 X_t\, dt\right)^p \leqslant 2\sqrt{10p}$$

since $\cos^2 \leqslant 1$. Similar application for the process $\sin X_t$ gives

$$\mathbb{E}\left[\left|\int_0^T \sin X_t\, dW_t\right|^{2p}\right] \leqslant 2\sqrt{10p}.$$

By Euler's formula, we can write $f(X_t) = \cos X_t + i \sin X_t$ and so

$$\int_0^T f(X_t)\, dW_t = \int_0^T \cos X_t\, dW_t + i\int_0^T \sin X_t\, dW_t.$$

Hence

$$\begin{aligned}
\mathbb{E}\left|\int_0^T f(X_t)\, dW_t\right|^{2p} &= \mathbb{E}\left[\left(\left|\int_0^T \cos X_t\, dW_t\right|^2 + \left|\int_0^T \sin X_t\, dW_t\right|^2\right)^p\right] \\
&\leqslant \mathbb{E}\left[2^p\left|\int_0^T \cos X_t\, dW_t\right|^{2p} + 2^p\left|\int_0^T \sin X_t\, dW_t\right|^{2p}\right] \\
&= 2^p\left(\mathbb{E}\left[\left|\int_0^T \cos X_t\, dW_t\right|^{2p}\right] + \mathbb{E}\left[\left|\int_0^T \sin X_t\, dW_t\right|^{2p}\right]\right) \leqslant 2^p 4\sqrt{10p}.
\end{aligned}$$

Moreover, as $\theta \in H_u$ we have by Lemma 3.2 that $\cos^2\theta \geqslant \frac{1}{2}$ and $\sin^2\theta \leqslant u^{-1}$. Hence

$$\left|\frac{\sigma i}{bi - \sigma^2/2}\right|^2 = \frac{\sigma^2}{b^2 + \sigma^4/4} \leqslant \frac{\sigma^2}{b^2} = \frac{4\pi^2 u^2 \sin^2\theta}{4\pi^2 u^2 \cos^2\theta} = \frac{\sin^2\theta}{\cos^2\theta} \leqslant 2u^{-1}.$$

Therefore,

$$\mathbb{E}\left|\int_0^T f(X_t)\, dt\right|^{2p} \leqslant 4\sqrt{10p}\, 4^p \theta^{2p} \leqslant 13p^{1/2}\, 4^p u^{-p}.$$

as required. □

Proof of Lemma 3.5. Fix $\varepsilon > 0$. Then for all $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ define the random variable

$$I(\boldsymbol{k}) := \left(\int_0^T Z_t\, dt\right) \cdot \chi_A(\boldsymbol{k}),$$

where χ_A is the indicator function on the set

$$A := \left\{\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\} : \theta \in H_u\right\}.$$

Note that $I(\boldsymbol{k})$ is well-defined and finite since $\left|\int_0^T Z_t\, dt\right| \leqslant 1$ by $|\exp(ix)| = 1$. Lemma 3.6 now yields for any $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ and $p \in \mathbb{N}$ that

$$\mathbb{E}|I(\boldsymbol{k})|^{2p} \leqslant 13p^{1/2}\, 4^p |\boldsymbol{k}|^{-p}$$

as when $\boldsymbol{k} \notin A$ we have $I(\boldsymbol{k}) \equiv 0$. Write $p_{\boldsymbol{k}} = \lfloor \log |\boldsymbol{k}| \rfloor$. Then

$$\mathbb{E} \sum_{\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}} |\boldsymbol{k}|^{-3} \frac{|I(\boldsymbol{k})|^{2p_k}}{13 p_{\boldsymbol{k}}^{1/2} 4^{p_k} |\boldsymbol{k}|^{-p_k}} \leqslant \sum_{\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}} |\boldsymbol{k}|^{-3} < \infty.$$

This means that the summands tend to 0 almost surely as $|\boldsymbol{k}| \to \infty$ and so we can find a random constant $C_\omega > 0$ such that for all $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ we have

$$|\boldsymbol{k}|^{-3} \frac{|I(\boldsymbol{k})|^{2p_k}}{13 p_{\boldsymbol{k}}^{1/2} 4^{p_k} |\boldsymbol{k}|^{-p_k}} \leqslant C_\omega.$$

Therefore, by possibly making C_ω bigger we obtain

$$|I(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2}.$$

This holds for each $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$, so by the definition of $I(\boldsymbol{k})$ we have whenever $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in H_u$ that

$$\left| \int_0^T Z_t \, dt \right| \leqslant C_\omega u^{-1/2}$$

as claimed. □

We are now in position to complete the proof of Lemma 3.3.

Proof of Lemma 3.3. Fix $\varepsilon > 0$. By the splitting

$$\hat{\mu}(\xi) = \int_0^1 Z_t \, dt = \int_0^T Z_t \, dt + \int_T^1 Z_t \, dt$$

and Lemmas 3.4 and 3.5, we have that almost surely there exists a constant $C_\omega > 0$ such that for all $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in H_u$ we have

$$|\hat{\mu}(\boldsymbol{k})| \leqslant \left| \int_0^T Z_t \, dt \right| + \left| \int_T^1 Z_t \, dt \right| \leqslant C_\omega |\boldsymbol{k}|^{-1/2}$$

as required. □

3C. *Vertical angles.* In this section we apply Kahane's work to obtain Fourier decay estimates when $\theta \in V_u$.

Lemma 3.7. *Fix $\varepsilon > 0$. Almost surely there exists a random constant $C_\omega > 0$ such that for any $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in V_u$ we have*

$$|\hat{\mu}(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2} \sqrt{\log |\boldsymbol{k}|}.$$

Let us discuss a few estimates obtained in [Kahane 1985b]. Let ν be the push-forward of Lebesgue measure on [0, 1] under the map $t \mapsto W_t$; that is, ν is the Brownian image of Lebesgue measure. Kahane established the following:

Theorem 3.8 [Kahane 1985b, page 255]. *Almost surely*

$$|\hat{\nu}(v)| \leqslant O(|v|^{-1} \sqrt{\log |v|}) \quad \textit{as } |v| \to \infty.$$

The key ingredient for the proof of Theorem 3.8 was based on establishing the following bound for the higher moments:

Lemma 3.9 [Kahane 1985b, page 254, estimate (2)]. *There exists a constant $C > 0$ such that for any $v \in \mathbb{R} \setminus \{0\}$ and any $p \in \mathbb{N}$ we have*

$$\mathbb{E}|\hat{\nu}(v)|^{2p} \leqslant C^p p^p |v|^{-2p}.$$

We can use Lemma 3.9 to give a bound on the higher moments in our setting, but with the price that the exponent will increase from $-2p$ to $-p$.

Lemma 3.10. *There exists a constant $C > 0$ such that for any $p \in \mathbb{N}$ and $\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\}$ with $\theta \in V_u$ the $(2p)$-th moment satisfies*

$$\mathbb{E}|\hat{\mu}(\xi)|^{2p} \leqslant C^p p^p |\xi|^{-p}.$$

Proof. Write $\boldsymbol{t} = (t_1, \dots, t_p) \in [0,1]^p$ and $d\boldsymbol{t}$ as the Lebesgue measure on $[0,1]^p$. Given $\boldsymbol{t}, \boldsymbol{s} \in [0,1]^p$, we define

$$\varphi(\boldsymbol{t}, \boldsymbol{s}) := \sum_{k=1}^{p} (t_k - s_k), \quad \psi(\boldsymbol{t}, \boldsymbol{s}) := \sum_{k=1}^{p} (W_{t_k} - W_{s_k}), \quad \text{and} \quad \Psi(\boldsymbol{t}, \boldsymbol{s}) := \mathbb{E}|\varphi(\boldsymbol{t}, \boldsymbol{s})|^2.$$

By the definition of μ_θ, μ and the Fourier transform, and using the fact that the multivariate process

$$X(\boldsymbol{t}, \boldsymbol{s}) := -2\pi \cos(\theta)\varphi(\boldsymbol{t}, \boldsymbol{s}) - 2\pi \sin(\theta)\psi(\boldsymbol{t}, \boldsymbol{s})$$

is Gaussian with mean $-2\pi\cos(\theta)\varphi(\boldsymbol{t}, \boldsymbol{s})$ and variance $4\pi^2 \sin^2(\theta)\Psi(\boldsymbol{t}, \boldsymbol{s})$, we have through Fubini's theorem and the formula for the characteristic function that

$$\begin{aligned}
\mathbb{E}|\hat{\mu}(\xi)|^{2p} &= \mathbb{E}\int_{[0,1]^p}\int_{[0,1]^p} \exp\bigl(-2\pi i u\bigl(\cos(\theta)\varphi(\boldsymbol{t}, \boldsymbol{s}) + \sin(\theta)\psi(\boldsymbol{t}, \boldsymbol{s})\bigr)\bigr)\, d\boldsymbol{t}\, d\boldsymbol{s} \\
&= \int_{[0,1]^p}\int_{[0,1]^p} \mathbb{E}\exp(iuX(\boldsymbol{t}, \boldsymbol{s}))\, d\boldsymbol{t}\, d\boldsymbol{s} \\
&= \int_{[0,1]^p}\int_{[0,1]^p} \exp\bigl(-2\pi i \cos(\theta)u\varphi(\boldsymbol{t}, \boldsymbol{s}) - 2\pi^2|u\sin(\theta)|^2\Psi(\boldsymbol{t}, \boldsymbol{s})\bigr)\, d\boldsymbol{t}\, d\boldsymbol{s}.
\end{aligned}$$

Thus by taking absolute values inside the integrals, and observing that $|\exp(ix)| = 1$ for any $x \in \mathbb{R}$, we obtain

$$\mathbb{E}|\hat{\mu}(\xi)|^{2p} \leqslant \int_{[0,1]^p}\int_{[0,1]^p} \exp(-2\pi^2|u\sin(\theta)|^2\Psi(\boldsymbol{t}, \boldsymbol{s}))\, d\boldsymbol{t}\, d\boldsymbol{s}. \tag{3-3}$$

On the other hand, by doing the expansion again for the Fourier transform $\hat{\nu}$ of the image measure ν at $v := u\sin(\theta) \in \mathbb{R} \setminus \{0\}$ we see that

$$\mathbb{E}|\hat{\nu}(v)|^{2p} = \mathbb{E}\int_{[0,1]^p}\int_{[0,1]^p} \exp(-2\pi i v\psi(\boldsymbol{t}, \boldsymbol{s}))\, d\boldsymbol{t}\, d\boldsymbol{s} = \int_{[0,1]^p}\int_{[0,1]^p} \exp(-2\pi^2 v^2\Psi(\boldsymbol{t}, \boldsymbol{s}))\, d\boldsymbol{t}\, d\boldsymbol{s},$$

which is equal to (3-3). Thus by Lemma 3.9 we have

$$\mathbb{E}|\hat{\mu}(\xi)|^{2p} \leqslant C^p p^p |v|^{-2p}.$$

Since $\theta \in V_u$ we have $|\sin\theta| \geqslant \min\left\{\frac{2}{\pi}u^{-1/2}, \frac{1}{\sqrt{2}}\right\}$. When $|\sin\theta| \geqslant \frac{1}{\sqrt{2}}$ we obtain

$$C^p p^p |v|^{-2p} \leqslant (2C)^p p^p u^{-2p} \leqslant (2C)^p p^p u^{-p}.$$

On the other hand, if $|\sin\theta| \geqslant \frac{2}{\pi}u^{-1/2}$ we have

$$C^p p^p |v|^{-2p} \leqslant C^p p^p (2u^{-1/2}/\pi)^{-2p} u^{-2p} \leqslant (C\pi^2/4)^p p^p u^{-p}. \qquad \square$$

Now we can complete the proof of Lemma 3.7 for vertical directions:

Proof of Lemma 3.7. Fix $\varepsilon > 0$. Then for all $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ define the random variable

$$F(\boldsymbol{k}) := \hat{\mu}(\boldsymbol{k})\chi_B(\boldsymbol{k}),$$

where

$$B := \left\{\xi = u(\cos\theta, \sin\theta) \in \mathbb{R}^2 \setminus \{0\} : \theta \in V_u\right\}.$$

Now $F(\boldsymbol{k})$ is a well-defined finite random variable as $|\hat{\mu}(\boldsymbol{k})| \leqslant 1$ for any $\boldsymbol{k}$. From Lemma 3.10 we obtain for any $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ and $p \in \mathbb{N}$ that

$$\mathbb{E}|F(\boldsymbol{k})|^{2p} \leqslant C^p p^p |\boldsymbol{k}|^{-p}.$$

Write $p_{\boldsymbol{k}} = \lfloor \log|\boldsymbol{k}| \rfloor$. Then

$$\mathbb{E} \sum_{\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}} |\boldsymbol{k}|^{-3} \frac{|F(\boldsymbol{k})|^{2p_{\boldsymbol{k}}}}{C^{p_{\boldsymbol{k}}} {p_{\boldsymbol{k}}}^{p_{\boldsymbol{k}}} |\boldsymbol{k}|^{-p_{\boldsymbol{k}}}} \leqslant \sum_{\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}} |\boldsymbol{k}|^{-3} < \infty.$$

This means that the summands tend to 0 almost surely as $|\boldsymbol{k}| \to \infty$ and so we can find a random constant $C_\omega > 0$ such that for all $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ we have

$$|\boldsymbol{k}|^{-3} \frac{|F(\boldsymbol{k})|^{2p_{\boldsymbol{k}}}}{C^{p_{\boldsymbol{k}}} {p_{\boldsymbol{k}}}^{p_{\boldsymbol{k}}} |\boldsymbol{k}|^{-p_{\boldsymbol{k}}}} \leqslant C_\omega.$$

Thus possibly making C_ω bigger, this yields

$$|F(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2} \sqrt{\log|\boldsymbol{k}|}.$$

Now this holds for each $\boldsymbol{k} \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$, so by the definition of $F(\boldsymbol{k})$ we have, whenever $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ with $\theta \in V_u$, that

$$|\hat{\mu}(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2} \sqrt{\log|\boldsymbol{k}|}$$

as claimed. $\square$

3D. ***From lattices to*** $\mathbb{R}^2$**.** We can now complete the proof of the main theorem. For this purpose, we need the following comparison lemma used by Kahane that allows us to pass from convergence on lattices for Fourier transforms to the whole space:

Lemma 3.11 [Kahane 1985b, Lemma 1, page 252]. *Suppose τ is a measure on $\mathbb{R}^2$ with support in $(-1, 1)^2$. Suppose $\varphi, \psi : (0, \infty) \to (0, \infty)$ are decreasing as $t \to \infty$ with the doubling properties*

$$\varphi(t/2) = O(\varphi(t)) \quad \textit{and} \quad \psi(t/2) = O(\psi(t)) \quad \textit{as } t \to \infty.$$

If the Fourier transform of τ along the integer lattice $\mathbb{Z}^2$ satisfies

$$|\hat{\tau}(\boldsymbol{n})| = O\left(\frac{\varphi(|\boldsymbol{n}|)}{\psi(|\boldsymbol{n}|)}\right) \quad \text{as } |\boldsymbol{n}| \to \infty,$$

then

$$|\hat{\tau}(\xi)| = O\left(\frac{\varphi(|\xi|)}{\psi(|\xi|)}\right) \quad \text{as } |\xi| \to \infty.$$

Proof of Theorem 1.3. Combining Lemmas 3.7 and 3.3 we have that for any $\varepsilon > 0$, almost surely, there exists some random constant $C_\omega > 0$ such that for any $\boldsymbol{k} = u(\cos\theta, \sin\theta) \in \varepsilon\mathbb{Z}^2 \setminus \{0\}$ we have

$$|\hat{\mu}(\boldsymbol{k})| \leqslant C_\omega |\boldsymbol{k}|^{-1/2}\sqrt{\log|\boldsymbol{k}|}. \tag{3-4}$$

Define a measure τ_ε on $\mathbb{R}^2$ such that

$$\hat{\tau}_\varepsilon(\xi) := \hat{\mu}(\varepsilon\xi), \quad \xi \in \mathbb{R}^2.$$

By the almost sure continuity of W_t, we have that there exists a random constant $M_\omega > 0$ such that the diameter of the support of μ is at most M_ω almost surely. Taking an intersection of the events that (3-4) holds for $\varepsilon = 1/n$ over all $n \in \mathbb{N}$ allows us to find a random $\varepsilon = \varepsilon_\omega > 0$ such that μ is supported on a set of diameter strictly less than $1/\varepsilon$ and (3-4) holds almost surely with this ε. This guarantees that the measure τ_ε is supported on $(-1, 1)^2$ and so applying Lemma 3.11 with the measure $\tau = \tau_\varepsilon$ and the maps $\varphi(t) := \sqrt{\log t}$ and $\psi(t) := t^{1/2}$ gives the claim. $\square$

Acknowledgements

We thank Tuomas Orponen for useful discussions during the preparation of this manuscript. We are also grateful to an anonymous referee for comments and suggestions which improved the focus of the paper. Finally, we thank The Hebrew University of Jerusalem and The University of Manchester for hosting us for research visits during the writing of this paper.

References

[Adler 1977] R. J. Adler, "Hausdorff dimension and Gaussian fields", *Ann. Probability* **5**:1 (1977), 145–151. MR Zbl

[Beffara 2008] V. Beffara, "The dimension of the SLE curves", *Ann. Probab.* **36**:4 (2008), 1421–1452. MR Zbl

[Bluhm 1998] C. Bluhm, "On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets", *Ark. Mat.* **36**:2 (1998), 307–316. MR Zbl

[Burkholder et al. 1972] D. L. Burkholder, B. J. Davis, and R. F. Gundy, "Integral inequalities for convex functions of operators on martingales", pp. 223–240 in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, II: Probability theory* (Berkeley, CA, 1970/1971), Univ. California Press, Berkeley, CA, 1972. MR Zbl

[Chan et al. 2016] V. Chan, I. Łaba, and M. Pramanik, "Finite configurations in sparse sets", *J. Anal. Math.* **128** (2016), 289–335. MR Zbl

[Davenport et al. 1963] H. Davenport, P. Erdős, and W. J. LeVeque, "On Weyl's criterion for uniform distribution", *Michigan Math. J.* **10** (1963), 311–314. MR Zbl

[Ekström et al. 2015] F. Ekström, T. Persson, and J. Schmeling, "On the Fourier dimension and a modification", *J. Fractal Geom.* **2**:3 (2015), 309–337. MR Zbl

[Fouché and Mukeru 2013] W. L. Fouché and S. Mukeru, “On the Fourier structure of the zero set of fractional Brownian motion”, *Statist. Probab. Lett.* **83**:2 (2013), 459–466. MR Zbl

[Fraser et al. 2014] J. M. Fraser, T. Orponen, and T. Sahlsten, “On Fourier analytic properties of graphs”, *Int. Math. Res. Not.* **2014**:10 (2014), 2730–2745. MR Zbl

[Jordan and Sahlsten 2016] T. Jordan and T. Sahlsten, “Fourier transforms of Gibbs measures for the Gauss map”, *Math. Ann.* **364**:3-4 (2016), 983–1023. MR Zbl

[Kahane 1966a] J.-P. Kahane, “Images browniennes des ensembles parfaits”, *C. R. Acad. Sci. Paris Sér. A-B* **263** (1966), A613–A615. MR Zbl

[Kahane 1966b] J.-P. Kahane, “Images d’ensembles parfaits par des séries de Fourier gaussiennes”, *C. R. Acad. Sci. Paris Sér. A-B* **263** (1966), A678–A681. MR Zbl

[Kahane 1985a] J.-P. Kahane, “Ensembles aléatoires et dimensions”, pp. 65–121 in *Recent progress in Fourier analysis* (El Escorial, 1983), edited by I. Peral and J. L. Rubio de Francia, North-Holland Math. Stud. **111**, North-Holland, Amsterdam, 1985. MR Zbl

[Kahane 1985b] J.-P. Kahane, *Some random series of functions*, 2nd ed., Cambridge Studies in Advanced Mathematics **5**, Cambridge University Press, 1985. MR Zbl

[Kahane 1993] J.-P. Kahane, “Fractals and random measures”, *Bull. Sci. Math.* **117**:1 (1993), 153–159. MR Zbl

[Karatzas and Shreve 1991] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Graduate Texts in Mathematics **113**, Springer, 1991. MR Zbl

[Kaufman 1980] R. Kaufman, “Continued fractions and Fourier transforms”, *Mathematika* **27**:2 (1980), 262–267. MR Zbl

[Kaufman 1981] R. Kaufman, “On the theorem of Jarník and Besicovitch”, *Acta Arith.* **39**:3 (1981), 265–267. MR Zbl

[Körner 2011] T. W. Körner, “Hausdorff and Fourier dimension”, *Studia Math.* **206**:1 (2011), 37–50. MR Zbl

[Łaba and Pramanik 2009] I. Łaba and M. Pramanik, “Arithmetic progressions in sets of fractional dimension”, *Geom. Funct. Anal.* **19**:2 (2009), 429–456. MR Zbl

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, 1995. MR Zbl

[Mattila 2015] P. Mattila, *Fourier analysis and Hausdorff dimension*, Cambridge Studies in Advanced Mathematics **150**, Cambridge University Press, 2015. MR Zbl

[McKean 1955] H. P. McKean, Jr., “Hausdorff–Besicovitch dimension of Brownian motion paths”, *Duke Math. J.* **22** (1955), 229–234. MR Zbl

[Peres and Sousi 2016] Y. Peres and P. Sousi, “Dimension of fractional Brownian motion with variable drift”, *Probab. Theory Related Fields* **165**:3-4 (2016), 771–794. MR Zbl

[Perkins 1981] E. Perkins, “The exact Hausdorff measure of the level sets of Brownian motion”, *Z. Wahrsch. Verw. Gebiete* **58**:3 (1981), 373–388. MR Zbl

[Peškir 1996] G. Peškir, “On the exponential Orlicz norms of stopped Brownian motion”, *Studia Math.* **117**:3 (1996), 253–273. MR Zbl

[Queffélec and Ramaré 2003] M. Queffélec and O. Ramaré, “Analyse de Fourier des fractions continues à quotients restreints”, *Enseign. Math.* (2) **49**:3-4 (2003), 335–356. MR Zbl

[Salem 1951] R. Salem, “On singular monotonic functions whose spectrum has a given Hausdorff dimension”, *Ark. Mat.* **1** (1951), 353–365. MR Zbl

[Shieh and Xiao 2006] N.-R. Shieh and Y. Xiao, “Images of Gaussian random fields: Salem sets and interior points”, *Studia Math.* **176**:1 (2006), 37–60. MR Zbl

[Shmerkin 2017] P. Shmerkin, “Salem sets with no arithmetic progressions”, *Int. Math. Res. Not.* **2017**:7 (2017), 1929–1941. MR

[Taylor 1953] S. J. Taylor, “The Hausdorff α-dimensional measure of Brownian paths in n-space”, *Proc. Cambridge Philos. Soc.* **49** (1953), 31–39. MR Zbl

[Taylor 1955] S. J. Taylor, “The α-dimensional measure of the graph and set of zeros of a Brownian path”, *Proc. Cambridge Philos. Soc.* **51** (1955), 265–274. MR Zbl

Received 28 Mar 2016. Revised 19 Jul 2017. Accepted 5 Sep 2017.

JONATHAN M. FRASER: jmf32@st-andrews.ac.uk
School of Mathematics and Statistics, University of St Andrews, St Andrews, United Kingdom

TUOMAS SAHLSTEN: tuomas.sahlsten@manchester.ac.uk
School of Mathematics, University of Manchester, Manchester, United Kingdom

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.133

NODAL GEOMETRY, HEAT DIFFUSION AND BROWNIAN MOTION

BOGDAN GEORGIEV AND MAYUKH MUKHERJEE

We use tools from n-dimensional Brownian motion in conjunction with the Feynman–Kac formulation of heat diffusion to study nodal geometry on a compact Riemannian manifold M. On one hand we extend a theorem of Lieb (1983) and prove that any Laplace nodal domain $\Omega_\lambda \subseteq M$ almost fully contains a ball of radius $\sim 1/\sqrt{\lambda_1(\Omega_\lambda)}$, and such a ball can be centred at any point of maximum of the Dirichlet ground state $\varphi_{\lambda_1(\Omega_\lambda)}$. This also gives a slight refinement of a result by Mangoubi (2008) concerning the inradius of nodal domains. On the other hand, we also prove that no nodal domain can be contained in a reasonably narrow tubular neighbourhood of unions of finitely many submanifolds inside M.

1. Introduction

We consider a compact n-dimensional smooth Riemannian manifold M, and the Laplacian (or the Laplace–Beltrami operator) $-\Delta$ on M. We use the analyst's sign convention; namely, $-\Delta$ is positive semidefinite. For an eigenvalue λ of $-\Delta$ and a corresponding eigenfunction φ_λ, recall that a nodal domain Ω_λ is a connected component of the complement of the nodal set $N_{\varphi_\lambda} := \{x \in M : \varphi_\lambda(x) = 0\}$. In this paper, we are interested in the asymptotic geometry of a nodal domain Ω_λ as $\lambda \to \infty$.

In this note we address the following two questions.

First, we start by discussing the problem of whether a nodal domain can be squeezed in a tubular neighbourhood around a certain subset $\Sigma \subseteq M$. A result of Steinerberger [2014, Theorem 2] states that for some constant $r_0 > 0$, a nodal domain Ω_λ cannot be contained in an $(r_0/\sqrt{\lambda})$-tubular neighbourhood of the hypersurface Σ, provided that Σ is sufficiently flat in the following sense: Σ must admit a unique metric projection in a wavelength (i.e., $\sim 1/\sqrt{\lambda}$) tubular neighbourhood. The proof involves the study of a heat process associated to the nodal domain, where one also uses estimates for Brownian motion and the Feynman–Kac formula.

We relax the conditions imposed on Σ. Our first result is a direct extension of [Steinerberger 2014, Theorem 2]. Before stating the result, we begin with the following definition:

Definition 1.1 (admissible collections). For each fixed eigenvalue λ, we consider a natural number $m_\lambda \in \mathbb{N}$ and a collection $\Sigma_\lambda := \bigcup_{i=1}^{m_\lambda} \Sigma_\lambda^i$, where Σ_λ^i is an embedded smooth submanifold (without boundary) of dimension k $(1 \leq k \leq n-1)$.

We call Σ_λ admissible up to a distance r if the following property is satisfied: for any $x \in M$ with $\operatorname{dist}(x, \Sigma_\lambda) \leq r$ there exists a unique index $1 \leq i_x(\lambda) \leq m_\lambda$ and a unique point $y \in \Sigma_\lambda^{i_x(\lambda)}$ realizing $\operatorname{dist}(x, \Sigma_\lambda)$, that is, $\operatorname{dist}(x, y) = \operatorname{dist}(x, \Sigma_\lambda)$.

MSC2010: primary 35P20, 53B20, 53Z05; secondary 35K05.
Keywords: Laplace eigenfunctions, nodal domains, Brownian motion.

We note that if Σ_λ consists of one submanifold which is admissible up to distance r, then Definition 1.1 means that r is smaller than the normal injectivity radius of Σ_λ. Moreover, if Σ_λ consists of more submanifolds, then these submanifolds must be disjoint and the distance between every two of them must be greater than r.

Let us also remark that, [Steinerberger 2014, Theorem 2] holds true when the hypersurface Σ is allowed to vary with respect to λ in a controlled way, which is made precise by Definition 1.1. With that clarification in place, our Theorem 1.2 is an extension of that result.

Theorem 1.2. *There is a constant r_0 depending only on (M, g) such that if a submanifold $\Sigma_\lambda \subset M$ is admissible up to distance $1/\sqrt{\lambda}$, then no nodal domain Ω_λ can be contained in an $(r_0/\sqrt{\lambda})$-tubular neighbourhood of Σ_λ.*

Further, it turns out that we can select Σ_λ to be a union of submanifolds of varying dimensions, having relaxed admissibility conditions.

Elaborating on this, we observe that getting entirely rid of the admissibility condition, as in Definition 1.1, allows situations where Σ_λ^i is dense in M, for example, $M = \mathbb{T}^2$ and Σ_λ^1 being a generic geodesic. By assuming Σ_λ^i is compact, we avoid such situations. Also, since we are considering unions of surfaces, the restriction of "unique projection" of nearby points, as in Definition 1.1, makes no sense anymore, and one can see that the approach of the proof of Theorem 1.2 does not work.

First, for ease of presentation, we adopt the following notation.

Definition 1.3. Given a compact subset K of M, let $\psi_K(t, x)$ denote the probability that a particle undergoing a Brownian motion starting at the point x will reach K within time t.

We now introduce the following relaxed notion of admissibility.

Definition 1.4 (α-admissible collections). Let $0 < \alpha < 1$ be a constant. For each fixed eigenvalue λ, we consider a natural number $m_\lambda \in \mathbb{N}$ and a collection $\Sigma_\lambda := \bigcup_{i=1}^{m_\lambda} \Sigma_\lambda^i$, where Σ_λ^i is a compact embedded smooth submanifold (without boundary) of dimension k_i, $(1 \le k_i \le n-1)$. Denote the respective tubular neighbourhoods by $N_\varepsilon(\Sigma_\lambda^i) := \{x \in M : \operatorname{dist}(x, \Sigma_\lambda^i) < \varepsilon\}$, and let $N_\varepsilon(\Sigma_\lambda) = \bigcup_{i=1}^{m_\lambda} N_\varepsilon(\Sigma_\lambda^i)$.

We say that the collection Σ_λ is α-admissible if for each sufficiently small $\varepsilon > 0$ and each $x \in N_\varepsilon(\Sigma_\lambda)$

$$\psi_{\partial B(x,2\varepsilon)\setminus N_\varepsilon(\Sigma_\lambda)}(4\varepsilon^2, x) \ge \alpha\, \psi_{\partial B(x,2\varepsilon)}(4\varepsilon^2, x). \tag{1}$$

Intuitively, using the above implicit formulation via Brownian motion hitting probabilities, we wish to ensure that $N_\varepsilon(\Sigma_\lambda)$ does not occupy too large a proportion of each $B(x, 2\varepsilon)$ for $x \in N_\varepsilon(\Sigma_\lambda)$; see Figure 2.

In other words, we allow the family Σ_λ to intersect, but the intersections should not be "too dense". To illustrate the idea, let us for simplicity assume that $M = \mathbb{R}^n$ and let us suppose that each member Σ_λ^i of the collection Σ_λ is a line passing through the origin. If the collection of these lines gets sufficiently close together or in other words "dense", then no matter how small an $\varepsilon > 0$ we take, the tubular neighbourhood $N_\varepsilon(\Sigma_\lambda)$ will contain the ball $B(0, 2\varepsilon)$. In particular, the left-hand side of (1) is vanishing and so, there is no $\alpha > 0$ for which the collection Σ_λ is α-admissible. Clearly, in the above example, replacing the lines Σ_λ^i by linear subspaces of varying dimensions will deliver a similar example of a collection, which is not α-admissible.

Having this intuition in mind, we have the following result.

Theorem 1.5. *Given an α-admissible collection Σ_λ, there exists a constant C, independent of λ, such that $N_{C/\sqrt{\lambda}}(\Sigma_\lambda)$ cannot fully contain a nodal domain Ω_λ.*

Theorem 1.5 gives a strong indication as to the "thickness" or general shape of a nodal domain in many situations of practical interest. For example, in dimension 2, numerics show nodal domains to look like a tubular neighbourhood of a tree. We also note that our proof of Theorem 1.5 reveals a bit more information, but for aesthetic reasons, we prefer to state the theorem this way. Heuristically, the proof reveals that the nodal domain Ω_λ is thicker at the points where the eigenfunction φ_λ attains its maximum, or at points where

$$\varphi_\lambda(x) \geq \beta \max_{y \in \Omega_\lambda} |\varphi_\lambda(y)|$$

for a fixed constant $\beta > 0$.

Second, we study the problem of how large a ball one may inscribe in a nodal domain Ω_λ at a point where the eigenfunction achieves extremal values on Ω_λ. We show:

Theorem 1.6. *Let $\dim M \geq 3$, $\varepsilon_0 > 0$ be fixed and $x_0 \in \Omega_\lambda$ be such that $|\varphi_\lambda(x_0)| = \max_{\Omega_\lambda} |\varphi_\lambda|$. There exists $r_0 = r_0(\varepsilon_0)$ such that*

$$\frac{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}) \cap \Omega_\lambda)}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} \geq 1 - \varepsilon_0. \tag{2}$$

A celebrated theorem of Lieb [1983] considers the case of a domain $\Omega \subset \mathbb{R}^n$ and states that there exists a point $x_0 \in \Omega$ where a ball of radius $C/\sqrt{\lambda_1(\Omega)}$ can almost be inscribed (in the sense of our Theorem 1.6). A further generalization was obtained in [Maz'ya and Shubin 2005] (see, in particular, Theorem 1.1 and Section 5.1 of that paper). However, the point x_0 was not specified. Physically, one expects that x_0 is close to the point where the first Dirichlet eigenfunction of Ω attains extremal values. This is in fact the essential statement of Theorem 1.6 above. Also, in this context, it is illuminating to compare the main theorem from [Croke and Derdziński 1987].

We reiterate that the proof of Theorem 1.6 uses estimates from [Grigor'yan and Saloff-Coste 2002], see (31), and a certain isocapacitary estimate, see (32), that work only in dimensions $n \geq 3$. As far as dimension $n = 2$ is concerned, it is known due to [Mangoubi 2008b, Theorem 1.2], see also [Hayman 1978], that any nodal domain has wavelength inradius; see further discussion on this at the beginning of Section 4.

As a corollary of Theorem 1.6, we derive the following:

Corollary 1.7. *Let M be a closed manifold of dimension $n \geq 3$, and $\Omega_\lambda \subseteq M$ be a nodal domain upon which the corresponding eigenfunction φ_λ is positive. Let x_0 be a point of maximum of φ_λ on Ω_λ. Then there exists a ball $B(x_0, C/\lambda^{\alpha(n)}) \subseteq \Omega_\lambda$ with $\alpha(n) = \frac{1}{4}(n-1) + \frac{1}{2n}$ and a constant $C = C(M, g)$.*

This recovers Theorem 1.5 of [Mangoubi 2008a], with the additional information that the ball of radius $C/\lambda^{\alpha(n)}$ is centred around the max point of the eigenfunction φ_λ (for more discussion on this, see Section 4). We also point out that using Theorem 1.6, the first author has established in [Georgiev 2016] using results from [Jakobson and Mangoubi 2009], the following inner radius bounds for real analytic manifolds:

Theorem 1.8 [Georgiev 2016]. *Let (M, g) be a real-analytic closed manifold of dimension at least* 3. *Let φ_λ be a Laplacian eigenfunction and Ω_λ be a nodal domain of φ_λ. Then, there exist constants c_1, c_2 depending only on (M, g) such that*

$$\frac{c_1}{\lambda} \leq \operatorname{inrad}(\Omega_\lambda) \leq \frac{c_2}{\sqrt{\lambda}}.$$

Moreover, if φ_λ is positive (*resp. negative*) *on Ω_λ, then a ball of this radius can be inscribed within a wavelength distance to a point where φ_λ achieves its maximum* (*resp. minimum*) *on Ω_λ.*

For another improvement of inner radius estimates in the smooth setting under certain conditional bounds on $\|\varphi_\lambda\|_{L^\infty(\Omega_\lambda)}$, see Theorem 1.7 of [Georgiev and Mukherjee 2016].

A few assorted remarks: as advertised, in Section 3 we address the problem of inscribing a nodal domain Ω_λ in a tubular neighbourhood around Σ. In this context, an interesting subcase one might also consider is Σ having conical singularities: at its singular points Σ looks locally like $\mathbb{R}^{n-1-k} \times \partial C^k$ for some $k = 1, \dots, n-1$, where ∂C^k denotes the boundary of a generalized cone, i.e., the cone generated by some open set $D \subseteq \mathbb{S}^{n-1}$.

In this situation a useful tool is an explicit heat kernel formula for generalized cones $C \subseteq \mathbb{R}^n$. One denotes the associated Dirichlet eigenfunctions and eigenvalues of the generating set D by m_j, l_j respectively. Using polar coordinates $x = \rho\theta$, $y = r\eta$, one has that the heat kernel of $P_C(t, x, y)$ of the generalized cone C is given by

$$P_C(t, x, y) = \frac{e^{-\frac{\rho^2+r^2}{2t}}}{t(\rho r)^{\frac{n}{2}+1}} \sum_{j=1}^{\infty} I_{\sqrt{l_j+(\frac{n}{2}-1)^2}}\left(\frac{\rho r}{t}\right) m_j(\theta)\, m_j(\eta), \tag{3}$$

where $I_\nu(z)$ denotes the modified Bessel function of order ν. For more on the formula (3) we refer to [Bañuelos and Smits 1997]. An even more general formula can be found in [Cheeger 1983].

The expression for $P_C(t, x, y)$ provides means for estimating $p_t(x)$ from below, as in Section 3. However, some features of the conical singularity (i.e., the eigenvalues and eigenfunctions l_j, m_j of the generating set D) enter explicitly in the estimate. Such considerations appear promising in discussing theorems of the following type, for example, and their higher-dimensional analogues; see also [Steinerberger 2014]:

Theorem 1.9 (Bers, Cheng). *Let $n = 2$. If $-\Delta u = \lambda u$, then any nodal set satisfies an interior cone condition with opening angle $\alpha \gtrsim \lambda^{-1/2}$.*

1A. ***Basic heuristics.*** We outline the main idea behind Theorems 1.2, 1.5 and 1.6.

First, one considers a point $x_0 \in \Omega_\lambda$ where the eigenfunction achieves a maximum on the nodal domain (without loss of generality we assume that the eigenfunction is positive on Ω_λ). One then considers the quantity $p(t, x_0)$, i.e., the probability that a Brownian motion started at x_0 escapes the nodal domain within time t.

The main strategy is to obtain two-sided bounds for $p(t, x_0)$.

On one hand, we have the Feynman–Kac formula (see Section 2A), which provides a straightforward upper bound only in terms of t (see (13) below).

On the other hand, depending on the context of the theorems above, we provide a lower bound for $p(t, x_0)$ in terms of some geometric data. To this end, we take advantage of various tools, some of which are: formulas for hitting probabilities of spheres and the parabolic scaling between the space and time variables, comparability of Brownian motions on manifolds with similar geometry (see Section 2B), bounds for hitting probabilities in terms of 2-capacity (see [Grigor'yan and Saloff-Coste 2002]), etc.

1B. ***Outline of the paper.*** In Section 2, we recall tools from n-dimensional Brownian motion and the Feynman–Kac formulation of heat diffusion, and discuss the parabolic scaling technique we referred to above. We include some background material on stochastic analysis on Riemannian manifolds, some of which (to our knowledge) is not widely known, but is important to our investigation. We also believe such results to be of independent interest to the community. Worthy of particular mention is Theorem 2.2, which roughly says that if the metric is perturbed slightly, hitting probabilities of compact sets by Brownian particles are also perturbed slightly. This allows us to apply Brownian motion formulae from $\mathbb{R}^n$ to compact manifolds, on small distance and time scales.

In Section 3, we begin by proving Theorem 1.2. As mentioned before, we then take the generalization one step further, by considering intersecting surfaces of different dimensions. Our main result in this direction is Theorem 1.5, which gives a quantitative lower bound on how "thin" or "narrow" a nodal domain can be.

In Section 4, we take up the investigation of inradius estimates of Ω_λ. As mentioned before, our main result in this direction is Theorem 1.6. We also establish Corollary 1.7.

2. Preliminaries: heat equation, Feynman–Kac and Bessel processes

2A. ***Feynman–Kac formula.*** We begin by stating a Feynman–Kac formula for open connected domains in compact manifolds for the heat equation with Dirichlet boundary conditions. Such formulas seem to be widely known in the community, but since we were unable to find out an explicit reference, we also indicate a line of proof.

Theorem 2.1. *Let M be a compact Riemannian manifold. For any open connected $\Omega \subset M$, $f \in L^2(\Omega)$, we have*

$$e^{t\Delta} f(x) = \mathbb{E}_x\big(f(\omega(t))\phi_\Omega(\omega, t)\big), \quad t > 0,\ x \in \Omega, \tag{4}$$

where $\omega(t)$ denotes an element of the probability space of Brownian motions starting at x, $\mathbb{E}_x$ is the expectation with regard to the measure on that probability space, and

$$\phi_\Omega(\omega, t) = \begin{cases} 1 & \text{if } \omega([0, t]) \subset \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

A proof of Theorem 2.1 can be constructed in three steps. First, one proves the corresponding statement when $\Omega = M$. This can be found, for example, in [Bär and Pfäffle 2011, Theorem 6.2]. One can then combine this with a barrier potential method to prove a corresponding statement for domains Ω with Lipschitz boundary. Lastly, the extension to domains with no regularity requirements on the boundary is achieved by a standard limiting argument. For details on the last two steps, see [Taylor 1996, Chapter 11, Section 3].

2B. ***Euclidean comparability of hitting probabilities.*** Implicit in many of our calculations is the following heuristic: if the metric is perturbed slightly, hitting probabilities of compact sets by Brownian particles are also perturbed slightly, provided one is looking at small distances r and at small time scales $t = O(r^2)$.

To describe the set up, let (M, g) be a compact Riemannian manifold and cover M by charts (U_k, ϕ_k) such that in these charts g is bi-Lipschitz to the Euclidean metric. Consider an open ball $B(p,r) \subset M$, where r is considered small, and in particular, smaller than the injectivity radius of M. Let $B(p,r)$ sit inside a chart (U, ϕ) and let $\phi(p) = q$ and $\phi(B(p,r)) = B(q,s) \subset \mathbb{R}^n$. Let K be a compact set inside $B(p,r)$ and let $K' := \phi(K) \subset B(q,s)$.

Now, let $\psi_K^M(T, p)$ denote the probability that a Brownian motion on (M, g) started at p and killed at a fixed time T hits K within time T. The probability $\psi_{K'}^e(t, q)$ is defined similarly for the standard Brownian motion in $\mathbb{R}^n$ started at q and killed at the same fixed time T. Now, we fix the time $T = cr^2$, where c is a constant. The following is the comparability result:

Theorem 2.2. *There exist constants c_1, c_2 depending only on c and M such that*

$$c_1 \psi_{K'}^e(T,q) \leq \psi_K^M(T,p) \leq c_2 \psi_{K'}^e(T,q). \tag{5}$$

The proof uses the concept of Martin capacity; see [Benjamini et al. 1995, Definition 2.1]:

Definition 2.3. Let Λ be a set and $\mathcal{B}$ a σ-field of subsets of Λ. Given a measurable function $F : \Lambda \times \Lambda \to [0, \infty]$ and a finite measure μ on $(\Lambda, \mathcal{B})$, the F-energy of μ is

$$I_F(\mu) = \int_\Lambda \int_\Lambda F(x,y)\, d\mu(x)\, d\mu(y).$$

The capacity of Λ in the kernel F is

$$\mathrm{Cap}_F(\Lambda) = \left[\inf_\mu I_F(\mu)\right]^{-1}, \tag{6}$$

where the infimum is over probability measures μ on $(\Lambda, \mathcal{B})$, and by convention, $\infty^{-1} = 0$.

Now we quote the following general result, which is Theorem 2.2 in [Benjamini et al. 1995].

Theorem 2.4. *Let $\{X_n\}$ be a transient Markov chain on the countable state space Y with initial state ρ and transition probabilities $p(x, y)$. For any subset Λ of Y, we have*

$$\tfrac{1}{2}\, \mathrm{Cap}_M(\Lambda) \leq \mathbb{P}_\rho\{\exists n \geq 0 : X_n \in \Lambda\} \leq \mathrm{Cap}_M(\Lambda), \tag{7}$$

where M is the Martin kernel $M(x,y) = G(x,y)/G(\rho,y)$, and $G(x,y)$ denotes the Green's function.

For the special case of Brownian motions, this reduces to (see Proposition 1.1 of [Benjamini et al. 1995] and Theorem 8.24 of [Mörters and Peres 2010]):

Theorem 2.5. *Let $\{B(t) : 0 \leq t \leq T\}$ be a transient Brownian motion in $\mathbb{R}^n$ starting from the point ρ, and $A \subset D$ be closed, where D is a bounded domain. Then,*

$$\tfrac{1}{2}\, \mathrm{Cap}_M(A) \leq \mathbb{P}_\rho\{B(t) \in A \ \textit{for some}\ 0 < t \leq T\} \leq \mathrm{Cap}_M(A). \tag{8}$$

An inspection of the proofs reveals that they go through with basically no changes on a compact Riemannian manifold M, when the Brownian motion is killed at a fixed time $T = cr^2$, and the Martin kernel $M(x, y)$ is defined as $G(x, y)/G(\rho, y)$, with $G(x, y)$ being the "cut-off" Green's function defined as follows: if $h_M(t, x, y)$ is the heat kernel of M,

$$G(x, y) := \int_0^T h_M(t, x, y)\, dt.$$

Now, to state it formally, in our setting, we have

Theorem 2.6.
$$\tfrac{1}{2}\operatorname{Cap}_M(K) \le \psi_K^M(T, p) \le \operatorname{Cap}_M(K). \tag{9}$$

Now, let $h_{\mathbb{R}^n}(t, x, y)$ denote the heat kernel on $\mathbb{R}^n$. To prove Theorem 2.2, it suffices to show that for $y \in K$, and $y' = \phi(y) \in K'$, we have constants C_1, C_2 (depending on c and M) such that

$$C_1 \int_0^T h_{\mathbb{R}^n}(t, q, y')\, dt \le \int_0^T h_M(t, p, y)\, dt \le C_2 \int_0^T h_{\mathbb{R}^n}(t, q, y')\, dt. \tag{10}$$

In other words, we need to demonstrate comparability of Green's functions "cut off" at time $T = cr^2$. Recall that we have the following Gaussian two-sided heat kernel bounds on a compact manifold (see, for example, Theorem 5.3.4 of [Hsu 2002] for the lower bound and Theorem 4 of [Cheng et al. 1981] for the upper bound, also (4.27) of [Grigor'yan and Saloff-Coste 2002]): for all $(t, p, y) \in (0, 1) \times M \times M$, and positive constants c_1, c_2, c_3, c_4 depending only on the geometry of M,

$$\frac{c_3}{t^{\frac{n}{2}}} e^{\frac{-c_1 d(p,y)^2}{4t}} \le h_M(t, p, y) \le \frac{c_4}{t^{\frac{n}{2}}} e^{\frac{-c_2 d(p,y)^2}{4t}},$$

where d denotes the distance function on M. Then, using the comparability of the distance function on M with the Euclidean distance function (which comes via metric comparability in local charts), for establishing (10), it suffices to observe that for any positive constant c_5, we have

$$\int_0^{cr^2} t^{-\frac{n}{2}} e^{-\frac{c_5 r^2}{4t}}\, dt = \frac{2^{n-2}}{c_5^{\frac{n}{2}-1}} \frac{1}{r^{n-2}} \Gamma\left(\frac{n}{2} - 1, \frac{c_5}{4c}\right),$$

where $\Gamma(s, x)$ is the (upper) incomplete Gamma function. Since r is a small constant chosen independently of λ, we observe that C_1, C_2 are constants in (10) depending only on $c, c_1, c_2, c_3, c_4, c_5, r$ and M, which finally proves (5).

Remark 2.7. Theorem 2.2 is implicit in [Steinerberger 2014], but it was not precisely stated or proved there. Since we are unable to find an explicit reference, here we have given a formal statement and indicated a proof. We believe that the statement of Theorem 2.2 will also be of independent interest for people interested in stochastic analysis on manifolds.

2C. ***Brownian motion on a manifold and Euclidean Bessel processes.*** Using the probabilistic formulation of the heat equation for the study of nodal geometry, we are largely inspired by the methods in [Steinerberger 2014]. Of course, such ideas have appeared in the literature before; for example, they are implicit in [Grieser and Jerison 1998]. Here we extend some ideas of Steinerberger with the help of tools from n-dimensional Brownian motion.

Given an open subset $V \subset M$, consider the solution $p_t(x)$ to the following diffusion process:

$$\begin{aligned}(\partial_t - \Delta) p_t(x) &= 0, \quad x \in V,\\ p_t(x) &= 1, \quad x \in \partial V,\\ p_0(x) &= 0, \quad x \in V.\end{aligned}$$

By the Feynman–Kac formula (see Section 2A), this diffusion process can be understood as the probability that a Brownian motion particle started in x will hit the boundary within time t. Now, fix an eigenfunction φ (corresponding to the eigenvalue λ) and a nodal domain Ω, so that $\varphi > 0$ on Ω without loss of generality. Calling Δ the Dirichlet Laplacian on Ω and setting $\Phi(t,x) := e^{t\Delta}\varphi(x)$, we see that Φ solves

$$\begin{aligned}(\partial_t - \Delta)\Phi(t,x) &= 0, && x \in \Omega,\\ \Phi(t,x) &= 0, && \text{on } \{\varphi = 0\},\\ \Phi(0,x) &= \varphi(x), && x \in \Omega.\end{aligned} \tag{11}$$

Using the Feynman–Kac formula given by Theorem 2.1, we have,

$$e^{t\Delta} f(x) = \mathbb{E}_x(f(\omega(t))\phi_\Omega(\omega,t)), \quad t > 0, \tag{12}$$

where $\omega(t)$ denotes an element of the probability space of Brownian motions starting at x, $\mathbb{E}_x$ is the expectation with regard to the measure on that probability space, and

$$\phi_\Omega(\omega,t) = \begin{cases} 1 & \text{if } \omega([0,t]) \subset \Omega,\\ 0 & \text{otherwise.}\end{cases}$$

Now, consider a nodal domain Ω corresponding to the eigenfunction φ, and consider the heat flow (11). Let $x_0 \in \Omega$ such that $\varphi(x_0) = \|\varphi\|_{L^\infty(\Omega)}$. We use the following upper bound derived in [Steinerberger 2014]:

$$\begin{aligned}\Phi(t,x) = e^{-\lambda t}\varphi(x) &= \mathbb{E}_x\big(\varphi(\omega(t))\phi_\Omega(\omega,t)\big)\\ &\le \|\varphi\|_{L^\infty(\Omega)}\mathbb{E}_x(\phi_\Omega(\omega,t)) = \|\varphi\|_{L^\infty(\Omega)}(1 - p_t(x)).\end{aligned} \tag{13}$$

Setting $t = \lambda^{-1}$ and $x = x_0$, we see that the probability of the Brownian motion starting at an extremal point x_0 leaving Ω within time λ^{-1} is $\le 1 - e^{-1}$. A rough interpretation is that maximal points x are situated deeply into the nodal domain. Using the notation introduced in the Introduction, the last derived upper estimate translates to $\psi_{M\setminus\Omega}(\lambda^{-1}, x) \le 1 - e^{-1}$.

Now, we consider an m-dimensional Brownian motion of a particle starting at the origin in $\mathbb{R}^m$, and calculate the probability of the particle hitting a sphere $\{x \in \mathbb{R}^m : \|x\| \le r\}$ of radius r within time t. By a well-known formula first derived in [Kent 1980], we see that such a probability is given as

$$\mathbb{P}\big(\sup_{0\le s\le t}\|B(s)\| \ge r\big) = 1 - \frac{1}{2^{\nu-1}\Gamma(\nu+1)}\sum_{k=1}^{\infty}\frac{j_{\nu,k}^{\nu-1}}{J_{\nu+1}(j_{\nu,k})}e^{-\frac{j_{\nu,k}^2 t}{2r^2}}, \quad \nu > -1, \tag{14}$$

where $\nu = (m-2)/2$ is the "order" of the Bessel process, J_ν is the Bessel function of the first kind of order ν, and $0 < j_{\nu,1} < j_{\nu,2} < \cdots$ is the sequence of positive zeros of J_ν.

Choose $x = x_0$, $t = \lambda^{-1}$, as before, and let $r = C^{1/2}\lambda^{-1/2}$, where C is a constant to be chosen later, independently of λ. Plugging this in (14) then reads as

$$\mathbb{P}\big(\sup_{0\le s\le\lambda^{-1}} \|B(s)\| \ge C\lambda^{-1/2}\big) = 1 - \frac{1}{2^{\nu-1}\Gamma(\nu+1)} \sum_{k=1}^{\infty} \frac{j_{\nu,k}^{\nu-1}}{J_{\nu+1}(j_{\nu,k})} e^{-\frac{j_{\nu,k}^2}{2C}}, \quad \nu > -1. \tag{15}$$

We need to make a few comments about the asymptotic behaviour of $j_{\nu,k}$ here. For notational convenience, we write $\alpha_k \sim \beta_k$ as $k \to \infty$ if we have $\alpha_k/\beta_k \to 1$ as $k \to \infty$. The asymptotic expansion

$$j_{\nu,k} = \big(k + \tfrac{1}{2}\nu + \tfrac{1}{4}\big)\pi + o(1) \quad \text{as } k \to \infty, \tag{16}$$

given in [Watson 1944, p. 506], tells us that $j_{\nu,k} \sim k\pi$. Also, from p. 505 of the same paper, we have

$$J_{\nu+1}(j_{\nu,k}) \sim (-1)^{k-1}\frac{\sqrt{2}}{\pi}\frac{1}{\sqrt{k}}. \tag{17}$$

These asymptotic estimates, in conjunction with (15), tell us that keeping ν bounded, and given a small $\eta > 0$, one can choose the constant C small enough (depending on η) such that

$$\mathbb{P}\big(\sup_{0\le s\le\lambda^{-1}} \|B(s)\| \ge C\lambda^{-1/2}\big) > 1 - \eta. \tag{18}$$

This estimate plays a role in Section 3. In this context, see also Proposition 5.1.4 of [Hsu 2002].

3. Admissibility conditions and intersecting surfaces

Proof of Theorem 1.2. If φ_λ attains its maximum within Ω_λ at x_0, we already know from (13) that

$$\psi_{M\setminus\Omega_\lambda}\left(\frac{t_0}{\lambda}, x_0\right) \le 1 - e^{-t_0}. \tag{19}$$

By the admissibility condition on Σ_λ we know that x_0 has a unique metric projection on one and only one $\Sigma_\lambda^{i_{x_0}}$ from the collection Σ_λ.

Now, suppose the result is not true. Choose R, t_0 small such that Theorem 2.2 applies. Choosing r_0 sufficiently smaller than R, we can find a λ such that Ω_λ is contained in an $(r_0/\sqrt{\lambda})$-tubular neighbourhood of Σ_λ, denoted by $N_{r_0\lambda^{-1/2}}(\Sigma_\lambda)$. From the remarks after Definition 1.1, it follows that $\Omega_\lambda \subseteq N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i_{x_0}})$.

We start a Brownian motion at x_0 and, roughly speaking, we see that locally the particle has freedom to wander in $n-k$ "bad directions", namely the directions normal to $\Sigma_\lambda^{i_{x_0}}$, before it hits $\partial\Omega_\lambda$. That means, we may consider a $(n-k)$-dimensional Brownian motion $B(t)$ starting at x_0; see Figure 1.

More formally, we choose a normal coordinate chart (U, ϕ) around x_0 adapted to $\Sigma_\lambda^{i_{x_0}}$, where the metric is comparable to the Euclidean metric. We have $\phi(\Sigma_\lambda^{i_{x_0}}) = \phi(U) \cap \{\mathbb{R}^k \times \{0\}^{n-k}\}$ and

$$\phi(N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i_{x_0}})) = \phi(U) \cap \left\{\mathbb{R}^k \times \left[-\frac{r_0}{\sqrt{\lambda}}, \frac{r_0}{\sqrt{\lambda}}\right]^{n-k}\right\}.$$

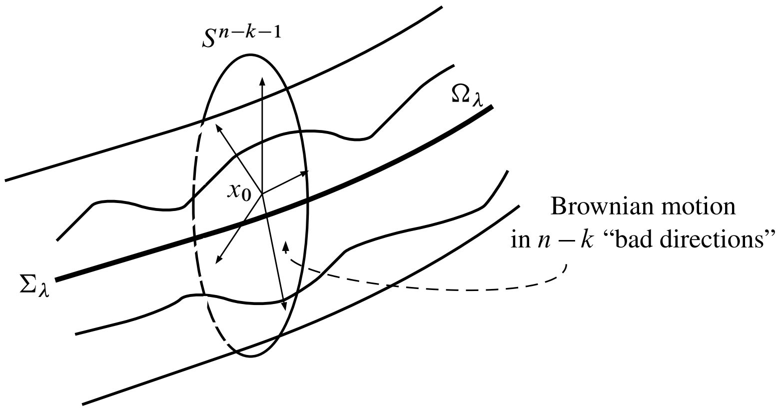

Figure 1. A Brownian motion at x_0.

We take a geodesic ball $B \subset U \subset M$ at x_0 of radius $R/\sqrt{\lambda}$. Using the hitting probability notation from Section 2 and monotonicity with respect to set inclusion we have

$$\psi_{M\setminus\Omega_\lambda}\left(\frac{t_0}{\lambda}, x_0\right) \geq \psi_{B\setminus\Omega_\lambda}\left(\frac{t_0}{\lambda}, x_0\right) \geq \psi_{B\setminus N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i x_0})}\left(\frac{t_0}{\lambda}, x_0\right), \tag{20}$$

and the comparability lemma implies that, if $c = t_0/R^2$, then there exists a constant C, depending on c and M, such that

$$\psi_{B\setminus N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i x_0})}\left(\frac{t_0}{\lambda}, x_0\right) \geq C\,\psi^e_{\phi(B\setminus N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i x_0}))}\left(\frac{t_0}{\lambda}, \phi(x_0)\right), \tag{21}$$

where ψ^e denotes the hitting probability in Euclidean space. We define $N^e_{r_0\lambda^{-1/2}} := \phi(N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i x_0}))$.

Let us consider the "solid cylinder" $S = B_{R/\sqrt{\lambda}} \times B_{r_0/\sqrt{\lambda}}$, a product of k and $(n-k)$-dimensional Euclidean balls centred at $\phi(x_0)$. S is clearly the largest cylinder contained in $N^e_{r_0\lambda^{-1/2}} \cap B$. We set $S = B_1 \times B_2$ for convenience. By monotonicity,

$$\psi^e_{\phi(B\setminus N_{r_0\lambda^{-1/2}}(\Sigma_\lambda^{i x_0}))}\left(\frac{t_0}{\lambda}, \phi(x_0)\right) \geq \psi^e_{B_1\times\partial B_2}\left(\frac{t_0}{\lambda}, \phi(x_0)\right). \tag{22}$$

If $B(t) = (B_1(t), \dots, B_n(t))$ is an n-dimensional Brownian motion, the components $B_i(t)$ are independent Brownian motions; see, for example, Chapter 2 of [Mörters and Peres 2010]. Denoting by $\mathcal{B}_k(t)$ and $\mathcal{B}_{n-k}(t)$ the projections of $B(t)$ onto the first k and last $n-k$ components respectively, it follows that

$$\begin{aligned}\psi^e_{B_1\times\partial B_2}\left(\frac{t_0}{\lambda}, \phi(x_0)\right) &\geq \mathbb{P}\left(\sup_{0\leq s\leq t_0\lambda^{-1}} \|\mathcal{B}_k(t)\| \leq \frac{R}{\sqrt{\lambda}}\right).\mathbb{P}\left(\sup_{0\leq s\leq t_0\lambda^{-1}} \|\mathcal{B}_{n-k}(t)\| \geq \frac{r_0}{\sqrt{\lambda}}\right)\\ &\geq c_k\mathbb{P}\left(\sup_{0\leq s\leq t_0\lambda^{-1}} \|\mathcal{B}_{n-k}(t)\| \geq \frac{r_0}{\sqrt{\lambda}}\right),\end{aligned}$$

where c_k is a constant depending on k and the ratio t_0/R^2, and can be calculated explicitly from (15).

Using the estimate in Section 2, we may take $r_0 \leq R$ sufficiently small so that

$$\mathbb{P}\left(\sup_{0 \leq s \leq t_0 \lambda^{-1}} \|\mathcal{B}_{n-k}(t)\| \geq \frac{r_0}{\sqrt{\lambda}} \right) > 1 - \varepsilon, \tag{23}$$

where ε is sufficiently small. Keeping $c = t_0/R^2$ and (hence) C fixed, we take t_0 small enough and $r_0 \leq R$ appropriately, so that (23) contradicts (20) and the fact that

$$\psi_{M \setminus \Omega_\lambda}(t_0 \lambda^{-1}, x) \leq 1 - e^{-t_0}. \qquad \square$$

Remark 3.1. Note that the constant r_0 above is independent of Σ_λ; in other words, the same constant r_0 will work for Theorem 1.2 as long as the surface is admissible up to a wavelength distance. Indeed, this results from the fact that r_0 depends only on the diffusion process associated to the Brownian motion, and is an inherent property of the manifold itself.

Now we address the generalizations of Theorem 1.2 for collections Σ_λ which are more complicated; namely, we assume Σ_λ is an α-admissible collection in the sense of Definition 1.4.

Proof of Theorem 1.5. By assumption, we have an α-admissible collection $\Sigma_\lambda := \bigcup_{i=1}^{m_\lambda} \Sigma_\lambda^i$.

Let us assume the contrary — if the statement is not true, we may select an arbitrarily small $r_0 > 0$ and find a corresponding inscribed nodal domain $\Omega_\lambda \subset N_{r_0 \lambda^{-1/2}}(\Sigma_\lambda)$.

As before, we choose a point $x_0 \in \Omega_\lambda$ such that

$$\varphi_\lambda(x_0) = \max_{x \in \Omega_\lambda} |\varphi_\lambda|.$$

Monotonicity of the hitting probability function $\psi_K(\cdot, \cdot)$ with respect to set inclusion in K, as well as the α-admissibility, imply that (see Figure 2)

$$\begin{aligned}
\psi_{M \setminus \Omega_\lambda}(t, x_0) &\geq \psi_{B(x_0, 2r_0 \lambda^{-1/2}) \setminus \Omega_\lambda}(t, x_0) \\
&\geq \psi_{B(x_0, 2r_0 \lambda^{-1/2}) \setminus N_{r_0 \lambda^{-1/2}}(\Sigma_\lambda)}(t, x_0) \\
&= \psi_{\partial(B(x_0, 2r_0 \lambda^{-1/2}) \setminus N_{r_0 \lambda^{-1/2}}(\Sigma_\lambda))}(t, x_0) \\
&\geq \psi_{\partial B(x_0, 2r_0 \lambda^{-1/2}) \setminus N_{r_0 \lambda^{-1/2}}(\Sigma_\lambda)}(t, x_0) \\
&\geq \alpha \psi_{\partial B(x_0, 2r_0 \lambda^{-1/2})}(t, x_0),
\end{aligned} \tag{24}$$

where we introduce the constant $\alpha > 0$ coming from the α-admissibility condition. Moreover, following Definition 1.4 of α-admissibility, in (24) we also assume that the radius $r_0/\sqrt{\lambda}$ is sufficiently small and that $t := t_0/\lambda$ with $t_0 := 4r_0^2$.

The latter estimate (24) implies, in particular, that

$$\frac{\psi_{M \setminus \Omega_\lambda}(t, x_0)}{\psi_{M \setminus B(x_0, 2r_0 \lambda^{-1/2})}(t, x_0)} = \frac{\psi_{M \setminus \Omega_\lambda}(t, x_0)}{\psi_{\partial B(x_0, 2r_0 \lambda^{-1/2})}(t, x_0)} \geq \alpha. \tag{25}$$

We now observe that by setting $t = t_0/\lambda$ we still have the freedom to choose t_0. We show that we can select t_0 such that (25) is violated. To this end we observe that the upper bound (19) along with (15) and

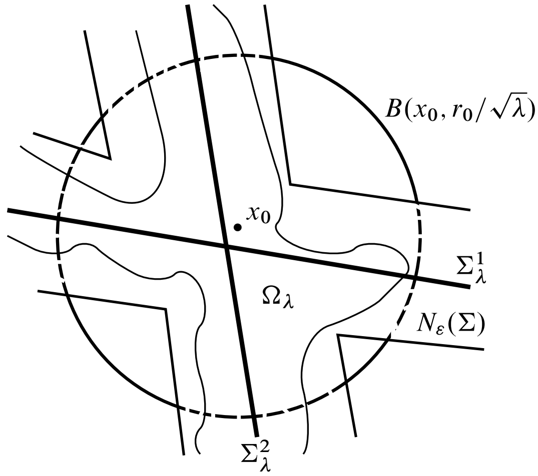

Figure 2. Nodal domain within a tubular neighbourhood of an admissible collection.

Theorem 2.2 give

$$\begin{aligned}\frac{\psi_{M\setminus\Omega_\lambda}(t_0/\lambda, x)}{\psi_{M\setminus B(x_0, 2r_0\lambda^{-1/2})}(t_0/\lambda, x)} &\lesssim (1-e^{-t_0})\biggl(1-\frac{1}{2^{\nu-1}\Gamma(\nu+1)}\sum_{k=1}^{\infty}\frac{j_{\nu,k}^{\nu-1}}{J_{\nu+1}(j_{\nu,k})}e^{-\frac{j_{\nu,k}^2 t_0}{2r_0^2}}\biggr)^{-1}\\ &= (1-e^{-t_0})\biggl(1-\frac{1}{2^{\nu-1}\Gamma(\nu+1)}\sum_{k=1}^{\infty}\frac{j_{\nu,k}^{\nu-1}}{J_{\nu+1}(j_{\nu,k})}e^{-2j_{\nu,k}^2}\biggr)^{-1}\\ &= (1-e^{-t_0})\widetilde{C}^{-1}. \end{aligned}\tag{26}$$

Now, we choose $t_0 = 4r_0^2$ small enough, so the last estimate yields a contradiction with (25). This proves the theorem. □

Remark 3.2. We wish to comment that in the above proof, it is not essential to look at the nodal domain only around the maximum point x_0. Given a predetermined positive constant β, choose a point $y \in \Omega_\lambda$ such that $\varphi_\lambda(y) \geq \beta\varphi_\lambda(x_0)$. Arguing similarly as in (13), we see that $\psi_{M\setminus\Omega_\lambda}(t, y) \leq 1-\beta e^{-t_0}$. Following the computations in (26), we get a constant r_0 (depending on β) such that $(1-\beta e^{-t_0})/\widetilde{C} < \alpha$, giving a contradiction. Also, it is clear that in Definitions 1.1 and 1.4, we do not actually need the submanifolds in the family Σ_λ to be smooth, and the proofs of Theorems 1.2 and 1.5 work with submanifolds of much lower regularity (for example, C^1 submanifolds).

4. Large ball at a max point

In this section we discuss the asymptotic thickness of nodal domains around extremal points of eigenfunctions. More precisely, let us consider a fixed nodal domain Ω_λ corresponding to the eigenfunction φ_λ. Let $x_0 \in \Omega_\lambda$ be such that

$$\varphi_\lambda(x_0) = \max_{x\in\Omega_\lambda} |\varphi_\lambda|. \tag{27}$$

In the case $\dim M = 2$, it was shown in Section 3 of [Mangoubi 2008b] that at such maximal points x_0 one can fully inscribe a large ball of wavelength radius (i.e $\sim 1/\sqrt{\lambda}$) into the nodal domain. In other words for Riemannian surfaces, one has that

$$\frac{C_1}{\sqrt{\lambda}} \leq \operatorname{inrad}(\Omega_\lambda) \leq \frac{C_2}{\sqrt{\lambda}}, \tag{28}$$

where C_i are constants depending only on M. Note that the proof for this case, as carried out in [Mangoubi 2008b] by following ideas in [Nazarov et al. 2005], makes use of essentially 2-dimensional tools (conformal coordinates and quasiconformality), which are not available in higher dimensions.

To our knowledge, in higher dimensions the sharpest known bounds on the inner radius of a nodal domain appear in [Mangoubi 2008a, Theorem 1.5] and state that

$$\frac{C_1}{\lambda^{\alpha(n)}} \leq \operatorname{inrad}(\Omega_\lambda) \leq \frac{C_2}{\sqrt{\lambda}}, \tag{29}$$

where $\alpha(n) := \frac{1}{4}(n-1) + \frac{1}{2n}$. A question of current investigation is whether the last lower bound on $\operatorname{inrad}(\Omega_\lambda)$ in higher dimensions is optimal.

Here we exploit heat equation and Brownian motion techniques to show that at least, one can expect to "almost" inscribe a large ball having radius to the order of $1/\sqrt{\lambda}$, in all dimensions. Now we prove Theorem 1.6:

Proof. We define $t' := t_0/\lambda$, and thus $\psi_{M\setminus\Omega_\lambda}(t', x) \leq 1 - e^{-t_0}$, where t_0 is a small constant to be chosen suitably later.

Now, choosing t_0 small enough, and using monotonicity, we have

$$\psi_{B(x_0, r_0\lambda^{-1/2})\setminus\Omega_\lambda}(t, x_0) < \psi_{M\setminus\Omega_\lambda}(t, x_0) < \varepsilon. \tag{30}$$

For convenience, let us define $E_{r_0} := B(x_0, r_0\lambda^{-1/2})\setminus\Omega_\lambda$, a relatively compact set. Observe that Theorem 2.2 applies to open balls and compact subsets contained in open balls. To adapt to the setting of Theorem 2.2, choose a number $r_0' < r_0$ such that $B(x_0, r_0'\lambda^{-1/2})$ satisfies

$$\frac{\operatorname{Vol}\big(B(x_0, r_0\lambda^{-1/2}) \setminus B(x_0, r_0'\lambda^{-1/2})\big)}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} < \varepsilon.$$

Call $E_{r_0'} := \overline{E_{r_0} \cap B(x_0, r_0'\lambda^{-1/2})}$. Observe that proving

$$\frac{\operatorname{Vol}(E_{r_0'})}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} < \varepsilon$$

will imply

$$\frac{\operatorname{Vol}(E_{r_0})}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} < 2\varepsilon,$$

which is what we want.

Now, we would like to compare the volumes of the two sets $E_{r_0'}$ and $B(x_0, r_0\lambda^{-1/2})$. Let $r = r_0/\sqrt{\lambda}$. Recall from [Grigor'yan and Saloff-Coste 2002, Remark 4.1] the following inequality:

$$c\frac{\operatorname{cap}(E_{r_0'})r^2}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))}e^{-C\frac{r^2}{t'}} \leq \psi_{E_{r_0}}(t', x_0) < \varepsilon, \tag{31}$$

where $\operatorname{cap}(K)$ denotes the 2-capacity of the set $K \subset M$, and $0 < t' < 2r^2$; see also equation (3.20) of [Grigor'yan and Saloff-Coste 2002]. Recall that the 2-capacity of a set $K \subset M$ is defined as

$$\operatorname{cap}(K) = \inf_{\substack{\eta|_K \equiv 1\\ \eta \in C^\infty(M)}} \int_M |\nabla\eta|^2 \, dM.$$

Formally, (31) holds on complete noncompact nonparabolic manifolds, which includes $\mathbb{R}^n$, $n \geq 3$. For bringing in our comparability result Theorem 2.2, we fix the ratio $t'/r^2 = \frac{1}{3}$, say, and then choose t_0 small enough that (30) still works. Now (31) applies, albeit with a new constant c as determined by the ratio t/r^2 and Theorem 2.2.

Now, to rewrite the capacity term in (31) in terms of volume, we bring in the following "isocapacitary inequality" [Maz'ya 2011, Section 2.2.3]:

$$\operatorname{cap}(E_{r_0}) \geq C' \operatorname{Vol}(E_{r_0})^{\frac{n-2}{n}}, \quad n \geq 3, \tag{32}$$

where C' is a constant depending only on the dimension n. We note that the isocapacitary inequality (in combination with a suitable Poincaré inequality) lies at the heart of the currently optimal inradius estimates, as derived in [Mangoubi 2008a].

Clearly, (31) and (32) together give

$$\left(\frac{\operatorname{Vol}(E_{r_0})}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))}\right)^{\frac{n-2}{n}} \lesssim \frac{\operatorname{cap}(E_{r_0})r^2}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} \lesssim \psi_{E_{r_0}}(t, x) < \varepsilon. \tag{33}$$

The last inequalities contain constants depending only on M, so by taking ε even smaller we can arrange

$$\frac{\operatorname{Vol}(E_{r_0})}{\operatorname{Vol}(B(x_0, r_0\lambda^{-1/2}))} < \varepsilon_0$$

for any initially given ε_0. □

Remark 4.1. We note that the heat equation method does not distinguish between a general domain and a nodal domain. This means that we cannot rule out the situation where $B(x_0, r_0/\sqrt{\lambda}) \setminus \Omega_\lambda$ is a collection of "sharp spikes" entering into $B(x_0, r_0/\sqrt{\lambda})$. Indeed the probability of a Brownian particle hitting a spike, no matter how "thin" it is, or how far from x_0 it is, is always nonzero, a fact related to the infinite speed of propagation of heat diffusion. This is consistent with the heuristic discussed in [Hayman 1978; Lieb 1983].

Now we establish Corollary 1.7. First, we recall the following result, which gives a bound on the asymmetry between the volumes of positivity and negativity sets:

Theorem 4.2 [Mangoubi 2008a]. *Let B be a geodesic ball, so that $\left(\frac{1}{2}B \cap \{\varphi_\lambda = 0\}\right) \neq \varnothing$ with $\frac{1}{2}B$ denoting the concentric ball of half radius. Then*

$$\frac{\mathrm{Vol}(\{\varphi_\lambda > 0\} \cap B)}{\mathrm{Vol}(B)} \geq \frac{C}{\lambda^{\frac{n-1}{2}}}. \tag{34}$$

Proof of Corollary 1.7. It suffices to combine the estimate (33) with (34).

Let $r := r_0/\sqrt{\lambda}$ be the radius of the largest inscribed ball in the nodal domain at x_0. Noting that $\{\varphi_\lambda < 0\} \subseteq E_{r_0}$ and combining Theorem 4.2 for $B_{x_0}(2r)$ with (33), we get

$$\left(\frac{C}{\lambda^{\frac{n-1}{2}}}\right)^{\frac{n-2}{n}} \leq \left(\frac{\mathrm{Vol}(E_{r_0})}{\mathrm{Vol}(B(x_0, r_0\lambda^{-1/2}))}\right)^{\frac{n-2}{n}} \leq 1 - e^{-\sqrt{1/3}r_0^2}. \tag{35}$$

Expanding the right-hand side in Taylor series and rearranging finishes the proof. □

Remark 4.3. An inspection of the proof of Theorem 1.6 reveals that one can take $\varepsilon = r_0^{2n/(n-2)}$. In other words, the relative volume of the error set E_{r_0} decays as $r_0^{2n/(n-2)}$ as $r_0 \to 0$. This is slightly better than the scaling prescribed by Corollary 2 of [Lieb 1983].

Remark 4.4. There is a sizable literature around optimizing the fundamental frequency of the complement of an obstacle inside a domain; for example, see [Harrell et al. 2001]. As an explicit special case, consider a convex domain $\Omega \subset \mathbb{R}^n$ and a small ball $B \subseteq \Omega$. The question is to find possible placements of translate $x + B$ inside Ω such that $\lambda_1(\Omega \setminus (x + B))$ is maximized. For certain applications of Theorem 1.6 towards such questions, we refer to [Georgiev and Mukherjee 2017].

Acknowledgements

It is a pleasure to thank Stefan Steinerberger for his detailed comments on a draft version of this paper, as well as Dan Mangoubi for his comments and remarks. The authors further thank Yuval Peres and Itai Benjamini for advice regarding Martin capacity, as well as Steve Zelditch for discussions on the Feynman–Kac formula and Brownian motion. Thanks are also due to the anonymous referee for a substantial improvement in the final presentation. Lastly, the authors would also like to thank Werner Ballmann, and gratefully acknowledge the Max Planck Institute for Mathematics, Bonn for providing ideal working conditions.

References

[Bañuelos and Smits 1997] R. Bañuelos and R. G. Smits, "Brownian motion in cones", *Probab. Theory Related Fields* **108**:3 (1997), 299–319. MR Zbl

[Bär and Pfäffle 2011] C. Bär and F. Pfäffle, "Wiener measures on Riemannian manifolds and the Feynman–Kac formula", *Mat. Contemp.* **40** (2011), 37–90. MR Zbl

[Benjamini et al. 1995] I. Benjamini, R. Pemantle, and Y. Peres, "Martin capacity for Markov chains", *Ann. Probab.* **23**:3 (1995), 1332–1346. MR Zbl

[Cheeger 1983] J. Cheeger, "Spectral geometry of singular Riemannian spaces", *J. Differential Geom.* **18**:4 (1983), 575–657. MR Zbl

[Cheng et al. 1981] S. Y. Cheng, P. Li, and S. T. Yau, “On the upper estimate of the heat kernel of a complete Riemannian manifold”, *Amer. J. Math.* **103**:5 (1981), 1021–1063. MR Zbl

[Croke and Derdziński 1987] C. B. Croke and A. Derdziński, “A lower bound for λ_1 on manifolds with boundary”, *Comment. Math. Helv.* **62**:1 (1987), 106–121. MR Zbl

[Georgiev 2016] B. Georgiev, “On the lower bound of the inner radius of nodal domains”, preprint, 2016. arXiv

[Georgiev and Mukherjee 2016] B. Georgiev and M. Mukherjee, “Some remarks on nodal geometry in the smooth setting”, preprint, 2016. arXiv

[Georgiev and Mukherjee 2017] B. Georgiev and M. Mukherjee, “On maximizing the fundamental frequency of the complement of an obstacle”, preprint, 2017. arXiv

[Grieser and Jerison 1998] D. Grieser and D. Jerison, “The size of the first eigenfunction of a convex planar domain”, *J. Amer. Math. Soc.* **11**:1 (1998), 41–72. MR Zbl

[Grigor’yan and Saloff-Coste 2002] A. Grigor’yan and L. Saloff-Coste, “Hitting probabilities for Brownian motion on Riemannian manifolds”, *J. Math. Pures Appl.* (9) **81**:2 (2002), 115–142. MR Zbl

[Harrell et al. 2001] E. M. Harrell, II, P. Kröger, and K. Kurata, “On the placement of an obstacle or a well so as to optimize the fundamental eigenvalue”, *SIAM J. Math. Anal.* **33**:1 (2001), 240–259. MR Zbl

[Hayman 1978] W. K. Hayman, “Some bounds for principal frequency”, *Applicable Anal.* **7**:3 (1978), 247–254. MR Zbl

[Hsu 2002] E. P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics **38**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Jakobson and Mangoubi 2009] D. Jakobson and D. Mangoubi, “Tubular neighborhoods of nodal sets and Diophantine approximation”, *Amer. J. Math.* **131**:4 (2009), 1109–1135. MR Zbl

[Kent 1980] J. T. Kent, “Eigenvalue expansions for diffusion hitting times”, *Z. Wahrsch. Verw. Gebiete* **52**:3 (1980), 309–319. MR Zbl

[Lieb 1983] E. H. Lieb, “On the lowest eigenvalue of the Laplacian for the intersection of two domains”, *Invent. Math.* **74**:3 (1983), 441–448. MR Zbl

[Mangoubi 2008a] D. Mangoubi, “Local asymmetry and the inner radius of nodal domains”, *Comm. Partial Differential Equations* **33**:7-9 (2008), 1611–1621. MR Zbl

[Mangoubi 2008b] D. Mangoubi, “On the inner radius of a nodal domain”, *Canad. Math. Bull.* **51**:2 (2008), 249–260. MR Zbl

[Maz’ya 2011] V. Maz’ya, *Sobolev spaces with applications to elliptic partial differential equations*, Grundlehren der Mathematischen Wissenschaften **342**, Springer, 2011. MR Zbl

[Maz’ya and Shubin 2005] V. Maz’ya and M. Shubin, “Can one see the fundamental frequency of a drum?”, *Lett. Math. Phys.* **74**:2 (2005), 135–151. MR Zbl

[Mörters and Peres 2010] P. Mörters and Y. Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics **30**, Cambridge University Press, 2010. MR Zbl

[Nazarov et al. 2005] F. Nazarov, L. Polterovich, and M. Sodin, “Sign and area in nodal geometry of Laplace eigenfunctions”, *Amer. J. Math.* **127**:4 (2005), 879–910. MR Zbl

[Steinerberger 2014] S. Steinerberger, “Lower bounds on nodal sets of eigenfunctions via the heat flow”, *Comm. Partial Differential Equations* **39**:12 (2014), 2240–2261. MR Zbl

[Taylor 1996] M. E. Taylor, *Partial differential equations, II: Qualitative studies of linear equations*, Applied Mathematical Sciences **116**, Springer, 1996. MR Zbl

[Watson 1944] G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press, 1944. MR Zbl

Received 11 Apr 2016. Revised 20 Jun 2017. Accepted 10 Aug 2017.

BOGDAN GEORGIEV: bogeor@mpim-bonn.mpg.de
Max Planck Institute for Mathematics, Bonn, Germany

MAYUKH MUKHERJEE: mukherjee@mpim-bonn.mpg.de
Max Planck Institute for Mathematics, Bonn, Germany

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.149

A NORMAL FORM À LA MOSER FOR DIFFEOMORPHISMS AND A GENERALIZATION OF RÜSSMANN'S TRANSLATED CURVE THEOREM TO HIGHER DIMENSIONS

JESSICA ELISA MASSETTI

We prove a discrete time analogue of Moser's normal form (1967) of real analytic perturbations of vector fields possessing an invariant, reducible, Diophantine torus; in the case of diffeomorphisms too, the persistence of such an invariant torus is a phenomenon of finite codimension. Under convenient nondegeneracy assumptions on the diffeomorphisms under study (a torsion property for example), this codimension can be reduced. As a by-product we obtain generalizations of Rüssmann's translated curve theorem in any dimension, by a technique of elimination of parameters.

1. Introduction and results

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, $a, b \in \mathbb{R}$, $a < b$, and consider the twist map

$$P : \mathbb{T} \times [a, b] \to \mathbb{T} \times \mathbb{R}, \quad (\theta, r) \mapsto (\theta + \alpha(r), r),$$

where $\alpha'(r) > 0$; this map preserves circles $r = r_0$, $r_0 \in [a, b]$, and rotates them by an angle which *increases* as r does (this is the twist property).

Moser [1962] proved that for any $r_0 \in (a, b)$ such that $\alpha(r_0)$ is Diophantine, if Q is an exact-area-preserving diffeomorphism sufficiently close to P, it has an invariant curve near $r = r_0$ on which the dynamics is conjugated to the rotation $\theta \mapsto \theta + \alpha(r_0)$.

Rüssmann [1970] generalized this fundamental result to nonconservative twist diffeomorphisms of the annulus; see also [Bost 1986; Yoccoz 1992]. He showed that the persistence of a Diophantine invariant

MSC2010: 37-XX, 37C05, 37J40.
Keywords: normal forms, Diophantine tori, KAM, counter terms, translated tori.

circle is a phenomenon of codimension 1; in general the invariant curve does not persist but it is translated in the normal direction. It is the "theorem of the translated curve" (see below for a precise statement).

As in Kolmogorov's theorem [1954], see also [Herman and Sergeraert 1971], the dynamics on the translated curve can be conjugated to the same initial Diophantine rotation because of the nondegeneracy (twist) of the map. Herman [1983] gave a proof of the translated curve theorem for diffeomorphisms with rotation number of constant type, then generalized Rüssmann's result in higher dimensions to diffeomorphisms of $\mathbb{T}^n \times \mathbb{R}$ ($\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$) close enough to the rotation $(\theta, r) \mapsto (\theta+\alpha, r)$, where α is a Diophantine vector, without assuming any twist hypothesis but introducing an external parameter in order to tune the frequency on the translated torus, yet breaking the dynamical conjugacy to the Diophantine rotation; see [Yoccoz 1992].

To our knowledge no further generalization in $\mathbb{T}^n \times \mathbb{R}^m$ of Rüssmann's theorem has been given so far.

The first purpose of this work is to prove a discrete-time analogue of Moser's normal form [1967] of real analytic perturbations of vector fields on $\mathbb{T}^n \times \mathbb{R}^m$ possessing a quasiperiodic Diophantine, reducible, invariant torus. The normal form will then be used to deduce "translated torus theorems" under convenient nondegeneracy assumptions. As a by-product, Rüssmann's classical theorem will be a particular case of small dimension. While Rüssmann and Herman consider smooth or finite differentiable diffeomorphisms, we focus here on the analytic category. Let us state the main results.

A normal form for diffeomorphisms. Let $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$ be the n-dimensional torus. Let V be the space of germs along $\mathbb{T}^n \times \{0\}$ in $\mathbb{T}^n \times \mathbb{R}^m = \{(\theta, r)\}$ of real analytic diffeomorphisms. Fix $\alpha \in \mathbb{R}^n$ and $A \in \mathrm{GL}_m(\mathbb{R})$, assuming that A is diagonalizable with (possibly complex) eigenvalues $a_1, \dots, a_m \in \mathbb{C}$.

Let $U(\alpha, A)$ be the affine subspace of V of diffeomorphisms of the form

$$P(\theta, r) = (\theta + \alpha + O(r), A \cdot r + O(r^2)), \tag{1-1}$$

where $O(r^k)$ are terms of order $\geq k$ in r which may depend on θ. For these diffeomorphisms, $\mathrm{T}^n_0 = \mathbb{T}^n \times \{0\}$ is an invariant, reducible, α-quasiperiodic torus whose normal dynamics at the first order is characterized by $a_1, \dots, a_m$. We will collectively refer to $\alpha_1, \dots, \alpha_n$ and $a_1, \dots, a_m$ as the *characteristic frequencies* or *characteristic numbers* of T^n_0.

Let now $a_1, \dots, a_q \in \mathbb{C}$ be the pairwise distinct eigenvalues of A. We will impose the following Diophantine conditions for some $\gamma > 0$ and $\tau \geq 1$:

$$\begin{aligned}
&\forall i = 1, \dots, q, && |a_i| = 1, && |k\cdot\alpha + \arg a_i - 2\pi l| \geq \frac{\gamma}{|k|^\tau} && \forall (k, l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z},\\
&\forall i, j = 1, \dots, q, && |a_i| = |a_j|, && |k\cdot\alpha + \arg a_i - \arg a_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau} && \forall (k, l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z},
\end{aligned} \tag{1-2}$$

where $\arg a_i \in [0, 2\pi[$ denotes the argument of the i-th eigenvalue $a_i = |a_i| e^{i \arg a_i}$.

Remark 1.1. Since A is in $\mathrm{GL}_m(\mathbb{R})$, the possible complex eigenvalues come in couples, and conditions (1-2) imply the classical Diophantine condition on α when $i = j$.

Let $\mathcal{G}$ be the space of germs of real analytic isomorphisms of $\mathbb{T}^n \times \mathbb{R}^m$ of the form

$$G(\theta, r) = (\varphi(\theta), R_0(\theta) + R_1(\theta) \cdot r), \tag{1-3}$$

where φ is a diffeomorphism of the torus fixing the origin and R_0, R_1 are functions defined on the torus $\mathbb{T}^n$ with values in $\mathbb{R}^m$ and $\mathrm{GL}_m(\mathbb{R})$ respectively and such that $\Pi_{\mathrm{Ker}(A-I)} R_0(0)=0$ and $\Pi_{\mathrm{Ker}[A,\cdot\,]}(R_1(0)-I)=0$, where I is the identity matrix in $\mathrm{Mat}_m(\mathbb{R})$ and Π is the projection on the indicated subspace.

Let us define the "correction map"

$$T_\lambda : \mathbb{T}^n \times \mathbb{R}^m \to \mathbb{T}^n \times \mathbb{R}^m, \quad (\theta, r) \mapsto (\beta+\theta, b+(I+B)\cdot r),$$

where $\beta \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $B \in \mathrm{Mat}_m(\mathbb{R})$ are such that

$$(A-I)\cdot b = 0, \quad [A, B] = 0. \tag{1-4}$$

We will refer to translating parameters $\lambda = (\beta, b + B\cdot r)$ as *corrections* or *counter terms*, and denote by Λ the space of such λ's:

$$\Lambda = \big\{\lambda = (\beta, b + B\cdot r) : (A-I)\cdot b = 0,\ [A, B] = 0\big\}.$$

Theorem A (normal form). *Let* (α, A) *satisfy the Diophantine condition* (1-2). *If* Q *is sufficiently close to* $P^0 \in U(\alpha, A)$, *there exists a unique triplet* $(G, P, \lambda) \in \mathcal{G} \times U(\alpha, A) \times \Lambda$ *close to* (id, P^0, 0), *such that*

$$Q = T_\lambda \circ G \circ P \circ G^{-1}.$$

In the neighborhood of (id, P^0, 0), the $\mathcal{G}$-orbit of all $P \in U(\alpha, A)$ has *finite* codimension. The proof is based on a relatively general inverse function theorem in analytic class (Theorem A.1 of the Appendix).

The idea of proving the finite codimension of a set of conjugacy classes of a diffeomorphism or of a vector field has been successfully exploited by many authors. Arnol'd [1961] first proved a normal form for diffeomorphisms of $\mathbb{T}^n$; this was followed by Moser's normal forms for vector fields [Moser 1966; 1967; Wagener 2010; Massetti 2015a; 2015b]. Among others, we recall the work of Calleja, Celletti and de la Llave [Calleja et al. 2013] on conformally symplectic systems, Chenciner's study [1985a; 1985b; 1988] on the bifurcation of elliptic fixed points, Herman's twisted conjugacy for Hamiltonians [Féjoz 2004; 2010] (a generalization of [Arnol'd 1961]) and the work of Eliasson, Fayad and Krikorian [Eliasson et al. 2015] around the stability of KAM tori.

This technique allows us to study the persistence of an invariant torus in two steps: first, prove a normal form that does not depend on any nondegeneracy hypothesis (but that contains the hard analysis); second, reduce or eliminate the (finite-dimensional) corrections by the usual implicit function theorem, using convenient nondegeneracy assumptions on the system under consideration. This second step was probably not deeply understood before the 80s [Sevryuk 1999].

A generalization of Rüssmann's theorem. From the normal form of Theorem A, we see that when $\lambda = 0$, $Q = G\circ P\circ G^{-1}$; the torus $G(\mathrm{T}^n_0)$ is invariant for Q and the first-order dynamics is given by $P \in U(\alpha, A)$. Conversely, whenever $\lambda = (\beta, b)$, the torus is *translated* and the α-quasiperiodic tangential dynamics is *twisted* by the correction β:

$$Q(\varphi(\theta), R_0(\theta)) = (\beta + \varphi(\theta+\alpha), b + R_0(\theta+\alpha)).$$

We will loosely say that the torus T^n_0

- persists up to twist-translation when $\lambda = (\beta, b)$,
- persists up to translation when $\lambda = (0, b)$.

We stress the fact that Theorem A not only gives the tangential dynamics to the torus, but also the normal one, of which Rüssmann's original statement is regardless:

Theorem (Rüssmann). *Let $\alpha \in \mathbb{R}$ be Diophantine and $P^0 : \mathbb{T} \times [-r_0, r_0] \to \mathbb{T} \times \mathbb{R}$ be of the form*

$$P^0(\theta, r) = \big(\theta + \alpha + t(r) + O(r^2),\, A^0 r + O(r^2)\big),$$

where $A^0 \in \mathbb{R} \setminus \{0\}$, $t(0) = 0$ and $t'(r) > 0$.

If Q is close enough to P^0, there exists a unique analytic curve $\gamma : \mathbb{T} \to \mathbb{R}$ close to $r = 0$, an analytic diffeomorphism φ of $\mathbb{T}$ fixing the origin close to the identity, and $b \in \mathbb{R}$ close to 0, such that

$$Q(\theta, \gamma(\theta)) = \big(\varphi \circ R_\alpha \circ \varphi^{-1}(\theta),\, b + \gamma(\varphi \circ R_\alpha \circ \varphi^{-1}(\theta))\big).$$

Note that $t(r)$ may depend on the angles as well. In the original statement, $A^0 = 1$; to consider this case with general A^0 does not add any difficulty to the proof.

We will generalize Rüssmann's theorem on $\mathbb{T}^n \times \mathbb{R}^n$. At the expense of losing control on the final normal dynamics and conjugating $T_\lambda^{-1} \circ Q$ to a diffeomorphism P whose invariant torus has a normal dynamics given by a different A, under convenient nondegeneracy conditions we can prove the existence of a twisted-translated or translated α-quasiperiodic Diophantine torus by application of the classical implicit function theorem in finite dimension. The following results will be proved in Section 5, where a more functional statement will be given (Theorems 5.1 and 5.4).

On $\mathbb{T}^n \times \mathbb{R}^n$, let $P \in U(\alpha, A)$, defined in expression (1-1), be such that A is invertible and has simple, real eigenvalues $a_1, \dots, a_n$. This hypothesis clearly implies that the only frequencies that can cause small divisors are the tangential ones $\alpha_1, \dots, \alpha_n$, so that we only need to require the standard Diophantine hypothesis on α.

Theorem B. *Let α be Diophantine and let $A \in \mathrm{GL}_n(\mathbb{R})$ have simple, real eigenvalues. If Q is sufficiently close to $P^0 \in U(\alpha, A)$, there exists A' close to A such that the torus T_0^n persists up to twist-translation and its final normal dynamics is given by A'.*

If, in addition, Q has a torsion property, we can prove the following theorem.

Theorem C. *Let α be Diophantine and let A be invertible with simple, real eigenvalues. Let also*

$$P^0(\theta, r) = \big(\theta + \alpha + p_1(\theta) \cdot r + O(r^2),\, A \cdot r + O(r^2)\big)$$

be such that

$$\det\left(\int_{\mathbb{T}^n} p_1(\theta)\, d\theta\right) \neq 0.$$

If Q is sufficiently close to P^0, there exists A' close to A such that the torus T_0^n persists up to translation and the final normal dynamics is given by A'.

The paper is organized as follows: in Sections 2–3 we introduce the normal form operator, define conjugacy spaces and present the difference equations that will be solved to linearize the dynamics on the perturbed torus; in Section 4 we will prove Theorem A, while in Section 5 we will prove Theorems B and C.

2. The normal form operator

We will show that the operator

$$\phi : \mathcal{G} \times U(\alpha, A) \times \Lambda \to V, \quad (G, P, \lambda) \mapsto T_\lambda \circ G \circ P \circ G^{-1},$$

is a local diffeomorphism (in the sense of scales of Banach spaces) in a neighborhood of (id, P^0, 0). Note that ϕ is formally defined on the whole space but $\phi(G, P, \lambda)$ is analytic in the neighborhood of T^n_0 only if G is close enough to the identity with respect to the width of analyticity of P. See the subsection on page 155.

Although the difficulty to overcome in the proof is rather standard for conjugacy problems of this kind (proving the fast convergence of a Newton-like scheme), the procedure relies on a relatively general inverse function theorem (Theorem A.1 of the Appendix), following a strategy different from Zehnder's [1975]. Both Zehnder's approach and ours rely on the fact that the fast convergence of the Newton scheme is somewhat independent of the internal structure of the variables.

Complex extensions. Let us extend the tori

$$\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n \quad \text{and} \quad \mathrm{T}^n_0 = \mathbb{T}^n \times \{0\} \subset \mathbb{T}^n \times \mathbb{R}^m,$$

as

$$\mathbb{T}^n_{\mathbb{C}} = \mathbb{C}^n/2\pi\mathbb{Z}^n \quad \text{and} \quad \mathrm{T}^n_{\mathbb{C}} = \mathbb{T}^n_{\mathbb{C}} \times \mathbb{C}^m$$

respectively, and consider the corresponding s-neighborhoods defined using ℓ^∞-balls (in the real normal bundle of the torus):

$$\mathbb{T}^n_s = \Big\{\theta \in \mathbb{T}^n_{\mathbb{C}} : \max_{1\le j\le n} |\mathrm{Im}\,\theta_j| \le s\Big\} \quad \text{and} \quad \mathrm{T}^n_s = \Big\{(\theta, r) \in \mathrm{T}^n_{\mathbb{C}} : |(\mathrm{Im}\,\theta, r)| \le s\Big\},$$

where $|(\mathrm{Im}\,\theta, r)| := \max\,(\max_{1\le j\le n}|\mathrm{Im}\,\theta_j|, \max_{1\le j\le m}|r_j|)$.

Let now $f : \mathrm{T}^n_s \to \mathbb{C}$ be holomorphic on the interior of T^n_s, continuous on T^n_s, and consider its Fourier expansion $f(\theta, r) = \sum_{k\in\mathbb{Z}^n} f_k(r)\, e^{i\,k\cdot\theta}$, where $k\cdot\theta = k_1\theta_1 + \cdots + k_n\theta_n$. In this context we introduce the so-called weighted norm:

$$|f|_s := \sum_{k\in\mathbb{Z}^n} |f_k|\, e^{|k|s}, \quad |k| = |k_1| + \cdots + |k_n|,$$

where $|f_k| = \sup_{|r|<s}|f_k(r)|$. Whenever $f : \mathrm{T}^n_s \to \mathbb{C}^n$, we have $|f|_s = \max_{1\le j\le n}(|f_j|_s)$, where f_j is the j-th component of $f(\theta, r)$.

It is a trivial fact that the classical sup-norm is bounded from above by the weighted norm:

$$\sup_{z\in\mathrm{T}^n_s} |f(z)| \le |f|_s$$

and that $|f|_s < +\infty$ whenever f is analytic on its domain, which necessarily contains some $\mathrm{T}^n_{s'}$ with $s' > s$. In addition, the following useful inequalities hold if f, g are analytic on $\mathrm{T}^n_{s'}$:

$$|f|_s \le |f|_{s'} \quad \text{for } 0 < s < s',$$

and

$$|fg|_{s'} \le |f|_{s'}|g|_{s'}.$$

Moreover, one can show that if f is analytic on $\mathrm{T}^n_{s+\sigma}$ and G is a diffeomorphism of the form (1-3) sufficiently close to the identity, then $|f \circ G|_s \le C_G |f|_{s+\sigma}$, where C_G is a positive constant depending on $|G - \mathrm{id}|_s$ (see Appendix C). For more details about the weighted norm, see for example [Meyer 1975; Chierchia 2003].

In general for complex extensions U_s and $V_{s'}$, we will denote by $\mathcal{A}(U_s, V_{s'})$ the set of holomorphic functions from U_s to $V_{s'}$ and by $\mathcal{A}(U_s)$, endowed with the s-weighted norm, the Banach space $\mathcal{A}(U_s, \mathbb{C})$.

Finally, let E and F be two Banach spaces:

– We indicate contractions with a dot "$\cdot$", with the convention that if $l_1, \dots, l_{k+p} \in E^*$ and $x_1, \dots, x_p \in E$,

$$(l_1 \otimes \cdots \otimes l_{k+p}) \cdot (x_1 \otimes \cdots \otimes x_p) = l_1 \otimes \cdots \otimes l_k \langle l_{k+1}, x_1 \rangle \cdots \langle l_{k+p}, x_p \rangle.$$

In particular, if $l \in E^*$, we simply write $l^n = l \otimes \cdots \otimes l$.

– If f is a differentiable map between two open sets of E and F, then $f'(x)$ is considered as a linear map belonging to $F \otimes E^*$, $f'(x) : \zeta \mapsto f'(x) \cdot \zeta$; the corresponding norm will be the standard operator norm

$$|f'(x)| = \sup_{\zeta \in E,\, |\zeta|_E = 1} |f'(x) \cdot \zeta|_F.$$

Spaces of conjugacies. – We consider the neighborhood of the identity $\mathcal{G}^\sigma_s$ in the space of germs of real holomorphic diffeomorphisms on T^n_s, defined by

$$|\varphi - \mathrm{id}|_s \le \sigma$$

and

$$|R_0 + (R_1 - I) \cdot r|_s \le \sigma,$$

where $\varphi(0) = 0$, and R_0 and R_1 satisfy $\Pi_{\ker(A-I)} R_0(0) = 0$ and $\Pi_{\ker([A,\cdot\,])}(R_1(0) - I) = 0$.

The tangent space at the identity $T_{\mathrm{id}}\mathcal{G}^\sigma_s$ consists of maps $\dot{G} \in \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^{n+m})$,

$$\dot{G}(\theta, r) = (\dot{\varphi}(\theta), \dot{R}_0(\theta) + \dot{R}_1(\theta) \cdot r),$$

where $\dot{\varphi} \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^n)$, $\dot{R}_0 \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ and $\dot{R}_1 \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$. We endow it with the norm

$$|\dot{G}|_s = \max_{1 \le j \le n+m} (|\dot{G}_j|_s).$$

– Let V_s be the subspace of $\mathcal{A}(\mathrm{T}^n_s, \mathbb{T}^n_{\mathbb{C}} \times \mathbb{C}^m)$ of diffeomorphisms

$$Q : (\theta, r) \mapsto (f(\theta, r), g(\theta, r)),$$

where $f \in \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^n)$, $g \in \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^m)$, endowed with the norm

$$|Q|_s = \max(|f|_s, |g|_s).$$

– Let $U_s(\alpha, A)$ be the affine subspace of V_s of those diffeomorphisms P of the form

$$P(\theta, r) = (\theta + \alpha + O(r), A \cdot r + O(r^2)).$$

We will indicate by p_i and P_i the coefficients of the order-i term in r, in the θ- and r-directions respectively.

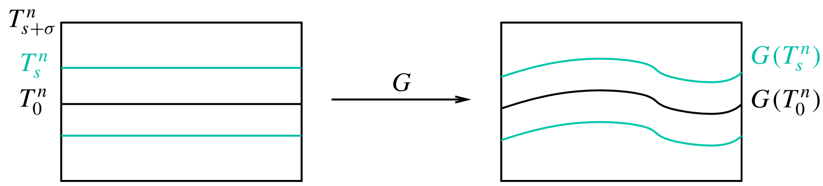

Figure 1. Deformed complex domain.

– If $G \in \mathcal{G}_s^\sigma$ and P is a diffeomorphism over $G(\mathrm{T}^n_s)$ we define the deformed norm

$$|P|_{G,s} := |P \circ G|_s,$$

depending on G; this is in order not to shrink artificially the domains of analyticity. See Figure 1. The problem, in a smooth context, may be solved without changing the domain, by using plateau functions.

The normal form operator. By Theorem B.1 and Corollary B.2 the operator

$$\begin{aligned}\phi : \mathcal{G}^{\sigma/n}_{s+\sigma} \times U_{s+\sigma}(\alpha, A) \times \Lambda \to V_s,\\ (G, P, \lambda) \mapsto T_\lambda \circ G \circ P \circ G^{-1},\end{aligned} \tag{2-1}$$

is now well defined. It would be more appropriate to write $\phi_{s,\sigma}$ but, since these operators commute with source and target spaces, we will refer to them simply as ϕ. We will always assume that $0 < s < s+\sigma < 1$ and $\sigma < s$.

3. Difference equations

We present here three lemmata that we will use in the following in order to linearize the tangent and the normal dynamics of the torus (see Section 4).

Let $\alpha \in \mathbb{R}^n$ and let $M \in \mathrm{GL}_m(\mathbb{R})$ have pairwise distinct eigenvalues $\mu_1, \dots, \mu_m$. We assume the following Diophantine conditions on α and M:

$$|k \cdot \alpha - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \qquad \forall k \in \mathbb{Z}^n \setminus \{0\},\ \forall l \in \mathbb{Z}, \tag{3-1}$$

$$|k \cdot \alpha - \arg \mu_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \qquad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z},\ \forall j = 1, \dots, m : |\mu_j| = 1, \tag{3-2}$$

$$|k \cdot \alpha + \arg \mu_i - \arg \mu_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \qquad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z},\ \forall i, j = 1, \dots, m : |\mu_i| = |\mu_j|, \tag{3-3}$$

$$\begin{cases} \big||\mu_i| - |\mu_j|\big| \geq \gamma & \forall i, j = 1, \dots, m,\ i \neq j : |\mu_i| \neq |\mu_j|, \\ \big|1 - |\mu_j|\big| \geq \gamma & \text{if } |\mu_j| \neq 1, \end{cases} \tag{3-4}$$

$$\begin{cases} |\mu_i - \mu_j| \geq \gamma & \forall i, j = 1, \dots, m,\ i \neq j : |\mu_i| = |\mu_j|, \\ |1 - \mu_j| \geq \gamma & \text{if } |\mu_j| = 1 \text{ and } \mu_j \neq 1, \end{cases} \tag{3-5}$$

$$\min_{1 \leq j \leq m} (|\mu_j|) \geq \gamma. \tag{3-6}$$

We first prove the following fundamental lemma, which is the heart of the proof of Theorem A and, more generally, of many stability results related to Diophantine rotations on the torus.

Lemma 3.1. *Let $\alpha \in \mathbb{R}^n$ be Diophantine in the sense of* (3-1) *and let $a, b \in \mathbb{C} \setminus \{0\}$.*

(1) *If $a = b$ and $|a| \geq \gamma$, for any $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$, there exists a unique f of zero average which is complex analytic on $\mathbb{T}^n_s$ and a unique $\lambda \in \mathbb{R}$ such that*

$$\lambda + af(\theta+\alpha) - af(\theta) = g(\theta), \quad \lambda = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} g\, d\theta,$$

satisfying

$$|f|_s \leq \frac{C}{\gamma^2\sigma^{\tau+n}}|g|_{s+\sigma},$$

where C is a constant depending only on n and τ.

(2) *Let $a \neq b$.*

(i) *If $|a| = |b|$ and*

$$\begin{cases} |a-b| \geq \gamma, \\ |a| \geq \gamma, \\ |k\cdot\alpha + \arg a - \arg b - 2\pi l| \geq \gamma/|k|^\tau \quad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z}, \end{cases} \tag{3-7}$$

for any $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$, there exists a unique f which is complex analytic on $\mathbb{T}^n_s$ such that

$$af(\theta+\alpha) - bf(\theta) = g(\theta), \tag{3-8}$$

satisfying

$$|f|_s \leq \frac{C}{\gamma^2\sigma^{\tau+n}}|g|_{s+\sigma},$$

where C is a constant depending only on n, τ.

(ii) *If $|a| \neq |b|$ and $\big||a| - |b|\big| \geq \gamma$, for any $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$, there exists a unique f which is complex analytic on $\mathbb{T}^n_{s+\sigma}$ such that*

$$af(\theta+\alpha) - bf(\theta) = g(\theta),$$

satisfying

$$|f|_{s+\sigma} \leq \gamma^{-1}|g|_{s+\sigma}.$$

Proof. (1) Developing in Fourier series the equation yields

$$\lambda + a\sum_k (e^{i\,k\cdot\alpha} - 1) f_k e^{i\,k\cdot\theta} = \sum_k g_k e^{i\,k\theta};$$

letting $\lambda = g_0$ we formally have

$$f(\theta) = \frac{1}{a}\sum_{k\neq 0} \frac{g_k}{e^{i\,k\,\alpha} - 1} e^{i\,k\theta}.$$

First note that the coefficients g_k decay exponentially, that is,

$$|g_k| = \left|\int_{\mathbb{T}^n} g(\theta) e^{-i\,k\cdot\theta}\,\frac{d\theta}{2\pi}\right| \leq |g|_{s+\sigma} e^{-|k|(s+\sigma)},$$

by deforming the path of integration to $\operatorname{Im}\theta_j = -\operatorname{sgn}(k_j)(s+\sigma)$.

Second, remark that for any $x, y \in \mathbb{R}^+$, $\varphi \in [0, 2\pi[$,

$$\begin{aligned} |x\,e^{i\varphi} - y|^2 &= (x-y)^2\cos^2\frac{\varphi}{2} + (x+y)^2\sin^2\frac{\varphi}{2} \\ &\geq (x+y)^2\sin^2\frac{\varphi}{2} = (x+y)^2\sin^2\frac{\varphi - 2\pi l}{2}, \end{aligned} \tag{3-9}$$

with $l \in \mathbb{Z}$. By choosing $l \in \mathbb{Z}$ such that $-\frac{\pi}{2} \leq \frac{1}{2}(\varphi - 2\pi l) \leq \frac{\pi}{2}$, we get

$$|x\,e^{i\varphi} - y| \geq \frac{2}{\pi}(x+y)\frac{|\varphi - 2\pi l|}{2}, \tag{3-10}$$

by the classical inequality $|\sin\delta| \geq \frac{2}{\pi}|\delta|$, whenever $-\frac{\pi}{2} \leq \delta \leq \frac{\pi}{2}$.

In our case $x = y = 1$, $\varphi = k\cdot\alpha$ and for all k, by choosing $l \in \mathbb{Z}$ such that $-\frac{\pi}{2} \leq \frac{1}{2}(k\cdot\alpha - 2\pi l) \leq \frac{\pi}{2}$, we get

$$|e^{i\,k\cdot\alpha} - 1| \geq \frac{4}{\pi}\frac{|k\cdot\alpha - 2\pi l|}{2} \geq \frac{2}{\pi}\frac{\gamma}{|k|^\tau},$$

by inequality (3-10) and the Diophantine condition (3-1).

We thus have

$$\begin{aligned} |f|_s &\leq \frac{\pi\,|g|_{s+\sigma}}{|a|\gamma}\sum_k |k|^\tau e^{-|k|\sigma} \leq \frac{\pi\,2^n|g|_{s+\sigma}}{|a|\gamma}\sum_{\ell\geq 1}\binom{\ell+n+1}{\ell}e^{-\ell\sigma}\ell^\tau \\ &\leq \frac{\pi\,4^n|g|_{s+\sigma}}{|a|\gamma(n-1)!}\sum_{\ell\geq 1}(n+\ell-1)^{n-1+\tau}e^{-\ell\sigma} \\ &\leq \frac{\pi\,4^n|g|_{s+\sigma}}{|a|\gamma(n-1)!}\int_1^\infty(\ell+n-1)^{n+\tau-1}e^{-(\ell-1)\sigma}\,d\ell. \end{aligned}$$

The integral is equal to

$$\sigma^{-\tau-n}e^{n\sigma}\int_{n\sigma}^\infty \ell^{\tau+n-1}e^{-\ell}\,d\ell < \sigma^{-\tau-n}e^{n\sigma}\int_0^\infty \ell^{\tau+n-1}e^{-\ell}\,d\ell = \sigma^{-\tau-n}e^{n\sigma}\Gamma(\tau+n).$$

Hence f, of zero average, is complex analytic on $\mathbb{T}^n_s$ and, since $|a| \geq \gamma$, it satisfies the claimed estimate.

(2i) Let $a = |a|e^{i\arg a}$ and $b = |b|e^{i\arg b}$ with the convention that $\arg z = \pi$ ($\arg z = 0$) if $z \in \mathbb{R}^-$ (if $z \in \mathbb{R}^+$).

The Fourier expansion gives

$$f_0 = \frac{g_0}{a-b}$$

and for all $k \neq 0$

$$f_k = \frac{g_k}{e^{i\arg b}\big(|a|e^{i(k\alpha+\arg a-\arg b)} - |b|\big)}e^{i\,k\cdot\theta}.$$

In order to bound the divisors we apply the same inequalities as in (3-9)–(3-10), with $\varphi = k\cdot\alpha + \arg a - \arg b$. Since $|a| = |b|$, by conditions (3-7) we proceed as in the proof of point (1) to get the stated estimate. In the case where a (or b) is real and $\arg a$ (or $\arg b$) is equal to π, we shall choose $\hat{l} = 2l-1$ (or $\hat{l} = 2l+1$) such that $-\frac{\pi}{2} \leq \frac{1}{2}(k\cdot\alpha + \arg a - \pi\hat{l}) \leq \frac{\pi}{2}$ to conclude the estimate as in (3-10).

(2ii) This follows directly from the triangular inequality.

We direct the reader interested to optimal estimates (with σ^τ instead of $\sigma^{\tau+n}$) to [Rüssmann 1976]. □

Let now $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$ have simple eigenvalues such that[1] $\mu_i \neq 1$ for all $i = 1, \dots, m$, and consider the operator

$$L_{1,M} : \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m) \to \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m), \quad f \mapsto f(\theta+\alpha) - M \cdot f(\theta).$$

Lemma 3.2 (relocating the torus). *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$, a diagonalizable matrix with simple eigenvalues distinct from* 1, *satisfy the Diophantine conditions* (3-1)–(3-2) *and* (3-4)–(3-6). *For every $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m)$, there exists a unique preimage $f \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ by $L_{1,A}$. Moreover, the following estimate holds*:

$$|f|_s \leq \frac{C_2}{\gamma^2} \frac{1}{\sigma^{n+\tau}} |g|_{s+\sigma},$$

where C_2 is a constant depending only on the dimension n and the exponent τ.

Proof. In the scalar case, $m = 1$ and $M = \mu \in \mathbb{R}$. By expanding both sides of $L_{1,M} f = g$, the formal preimage is given by

$$f_k = \frac{g_k}{e^{i\,k\,\alpha} - \mu}$$

and the proof is recovered from Lemma 3.1(2ii). The diagonal case follows readily by working componentwise and taking into account condition (3-4).

Finally, if M is diagonalizable, let $P \in \mathrm{GL}_m(\mathbb{C})$ be the diagonalizing matrix such that $PMP^{-1} = \mathrm{diag}(\mu_1, \dots, \mu_m)$, $\mu_i \in \mathbb{C}$. By left multiplying both sides of $f(\theta+\alpha) - M \cdot f(\theta) = g$ by P, we get

$$\tilde{f}(\theta+\alpha) - PMP^{-1} \tilde{f}(\theta) = \tilde{g},$$

where we have set $\tilde{g} = Pg$ and $\tilde{f} = Pf$. By Lemma 3.1(2) and the Diophantine conditions (3-1)–(3-2) and (3-5)–(3-6), $\tilde{f}$ satisfies the wanted estimates, and $f = P^{-1}\tilde{f}$. □

Finally, consider a holomorphic function F on $\mathbb{T}^n_{s+\sigma}$ with values in $\mathrm{Mat}_m(\mathbb{C})$ and define the operator

$$\begin{aligned} L_{2,M} : \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C})) &\to \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C})), \\ F &\mapsto F(\theta+\alpha) \cdot M - M \cdot F(\theta). \end{aligned}$$

Lemma 3.3 (straighten the first order dynamics). *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$, a diagonalizable matrix with simple eigenvalues distinct from* 1, *satisfy the Diophantine conditions* (3-1) *and* (3-3)–(3-6). *For every $G \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C}))$ such that $\int_{\mathbb{T}^n} G^i_i/(2\pi)^n \, d\theta = 0$ there exists a unique $F \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$, having zero-average diagonal elements, such that the matrix equation*

$$F(\theta+\alpha) \cdot M - M \cdot F(\theta) = G(\theta)$$

is satisfied; *moreover*, *the following estimate holds*:

$$|F|_s \leq \frac{C_3}{\gamma^2} \frac{1}{\sigma^{n+\tau}} |G|_{s+\sigma},$$

where C_3 is a constant depending only on the dimension n and the exponent τ.

[1] In order not to burden the following statements, we suppose that M has simple spectrum and 1 does not belong to it. Just note that in the general case, one should introduce the correction λ meant to absorb the average of the given term in the homological equations when it is the case, as in Lemma 3.1(1); cf. conditions (1-4).

Proof. Let $M = \operatorname{diag}(\mu_1, \dots, \mu_m) \in \mathbb{R}^m$ and $F \in \operatorname{Mat}_m(\mathbb{C})$ be given; expanding $L_{2,M}F = G$ we get m equations of the form

$$\mu_j(F_j^j(\theta+\alpha) - F_j^j(\theta)) = G_j^j, \quad j = 1, \dots, m,$$

and $m^2 - m$ equations of the form

$$\mu_j F_j^i(\theta+\alpha) - \mu_i F_j^i(\theta) = G_j^i(\theta), \quad \forall i \neq j,\ i, j = 1, \dots, m,$$

where we denoted by F_j^i the element corresponding to the i-th line and j-th column of the matrix $F(\theta)$. Taking into account the Diophantine conditions (3-1)–(3-4), the thesis follows from the same computations as Lemma 3.1(1) for the m-diagonal equations and point (2ii) for the (m^2-m)-out diagonal ones.

Finally, to recover the general case, we consider the transition matrix $P \in \mathrm{GL}_m(\mathbb{C})$ such that $PMP^{-1} = \operatorname{diag}(\mu_1, \dots, \mu_m)$, $\mu_i \in \mathbb{C}$, and the equation

$$(PF(\theta+\alpha)P^{-1}PMP^{-1}) - PMP^{-1}PF(\theta)P^{-1} = PGP^{-1};$$

letting $\widetilde{F} = PFP^{-1}$ and $\widetilde{G} = PGP^{-1}$, the equation is of the previous kind and by the Diophantine conditions (3-1) and (3-3)–(3-6), $\widetilde{F}$ satisfies the wanted estimates, and $F = P^{-1}\widetilde{F}P$. □

Remark 3.4. The real analytic character of the solutions in Lemmata 3.2 and 3.3 follows from their uniqueness and the fact that the matrix M has real entries.

4. Inversion of the operator ϕ

The following theorem represents the main result of this first part, from which the normal form theorem, Theorem A, follows.

Let us fix $P^0 \in U_s(\alpha, A)$ and note $V_s^\sigma = \{Q \in V_s : |Q - P^0|_s < \sigma\}$, the ball of radius σ centered at P^0.

Theorem 4.1. *The operator ϕ is a local diffeomorphism in the sense that for any $0 < \eta < s < s+\sigma < 1$ there exists $\varepsilon > 0$ and a unique C^∞-map ψ,*

$$\psi : V_{s+\sigma}^\varepsilon \to \mathcal{G}_s^\eta \times U_s(\alpha, A) \times \Lambda,$$

such that $\phi \circ \psi = \mathrm{id}$. Moreover, ψ is Whitney-smooth with respect to (α, A).

This result will follow from the inverse function theorem, Theorem A.1, and regularity propositions, Propositions A.2–A.4.

In order to solve locally $\phi(x) = y$, we use the remarkable idea of Kolmogorov and find the solution by composing infinitely many times the operator

$$x = (g, u, \lambda) \mapsto x + \phi'^{-1}(x) \cdot (y - \phi(x))$$

on extensions $\mathrm{T}_{s+\sigma}^n$ of shrinking width.

At each step of the induction, it is necessary that $\phi'^{-1}(x)$ exists at an unknown x (not only at x_0) in a whole neighborhood of x_0 and that ϕ'^{-1} and ϕ'' satisfy a suitable estimate, in order to control the convergence of the iterates.

The main step is to check the existence of a right inverse for

$$\phi'(G, P, \lambda) : T_G \mathcal{G}^{\sigma/n}_{s+\sigma} \times \vec{U}_{s+\sigma} \times \Lambda \to V_{G,s}$$

if G is close to the identity. We denote by $\vec{U}$ the vector space directing $U(\alpha, A)$.

Proposition 4.2. *If* (G, P, λ) *is close enough to* $(\mathrm{id}, P^0, 0)$ *for all* $\delta Q \in V_{G,s+\sigma} = G^* \mathcal{A}(\mathrm{T}^n_{s+\sigma}, \mathbb{C}^{n+m})$, *there exists a unique triplet* $(\delta G, \delta P, \delta\lambda) \in T_G \mathcal{G}_s \times \vec{U}_s \times \Lambda$ *such that*

$$\phi'(G, P, \lambda) \cdot (\delta G, \delta P, \delta\lambda) = \delta Q. \tag{4-1}$$

Moreover, we have the estimate

$$\max(|\delta G|_s, |\delta P|_s, |\delta\lambda|) \le \frac{C'}{\sigma^{\tau'}} |\delta Q|_{G,s+\sigma}, \tag{4-2}$$

where C' *is a constant possibly depending on* $|((G - \mathrm{id}), P - (\theta + \alpha, A \cdot r))|_{s+\sigma}$.

Proof. Let a vector field $\delta Q \in V_{G,s+\sigma}$ be given. Differentiating with respect to $x = (G, P, \lambda)$, we have

$$\delta(T_\lambda \circ G \circ P \circ G^{-1}) = T_{\delta\lambda} \circ (G \circ P \circ G^{-1}) + T'_\lambda \circ (G \circ P \circ G^{-1}) \cdot \delta(G \circ P \circ G^{-1});$$

hence

$$M \cdot (\delta G \circ P + G' \circ P \cdot \delta P - G' \circ P \cdot P' \cdot G'^{-1} \cdot \delta G) \circ G^{-1} = \delta Q - T_{\delta\lambda} \circ (G \circ P \circ G^{-1}),$$

where $M = \left(\begin{smallmatrix} I & 0 \\ 0 & I+B \end{smallmatrix}\right)$.

The data is δQ, while the unknowns are the "tangent vectors" $\delta P \in O(r) \times O(r^2)$, δG (geometrically, a vector field along G) and $\delta\lambda \in \Lambda$.

Precomposing by G, we get the equivalent equation between germs along the standard torus T^n_0 (as opposed to $G(\mathrm{T}^n_0)$):

$$M \cdot (\delta G \circ P + G' \circ P \cdot \delta P - G' \circ P \cdot P' \cdot G'^{-1} \cdot \delta G) = \delta Q \circ G - T_{\delta\lambda} \circ G \circ P;$$

multiplying both sides by $(G'^{-1} \circ P) M^{-1}$, we finally obtain

$$\dot{G} \circ P - P' \cdot \dot{G} + \delta P = G'^{-1} \circ P \cdot M^{-1} \delta Q \circ G + G'^{-1} \circ P \cdot M^{-1} T_{\delta\lambda} \circ G \circ P, \tag{4-3}$$

where $\dot{G} = G'^{-1} \cdot \delta G$.

Note that the term containing $T_{\delta\lambda}$ is not constant; expanding along $r = 0$, it reads as

$$T_{\dot{\lambda}} = G'^{-1} \circ P \cdot M^{-1} \cdot T_{\delta\lambda} \circ G \circ P = (\dot{\beta} + O(r), \dot{b} + \dot{B} \cdot r + O(r^2)).$$

The vector field $\dot{G}$ (geometrically, a germ along T^n_0 of tangent vector fields) reads as

$$\dot{G}(\theta, r) = (\dot{\varphi}(\theta), \dot{R}_0(\theta) + \dot{R}_1(\theta) \cdot r).$$

The problem is now: G, λ, P, Q being given, find $\dot{G}$, δP and $\dot{\lambda}$, and hence $\delta\lambda$ and δG.

We are interested in solving the equation up to the 0-order in r in the θ-direction, and up to the first order in r in the action direction; hence we consider the Taylor expansions along T^n_0 up to the needed order.

We remark that since $\delta P = (O(r), O(r^2))$, it will not intervene in the cohomological equations given out by (4-3), but will be uniquely determined by identification of the reminders.

Let us proceed to solve (4-3); taking its jet at the wanted order, it splits into the following three equations:

$$\begin{aligned} \dot\varphi(\theta+\alpha)-\dot\varphi(\theta)+p_1\cdot \dot R_0 &= \dot q_0+\dot\beta,\\ \dot R_0(\theta+\alpha)-A\cdot \dot R_0(\theta) &= \dot Q_0+\dot b,\\ \dot R_1(\theta+\alpha)\cdot A - A\cdot \dot R_1(\theta) &= \dot Q_1-(2P_2\cdot \dot R_0+\dot R_0'(\theta+\alpha)\cdot p_1)+\dot B. \end{aligned}\tag{4-4}$$

The first equation is the one straightening the tangential dynamics, while the second and the third ones are meant to relocate the torus and straighten the normal dynamics.

For the moment we solve the equations "modulo $\dot\lambda$"; eventually $\delta\lambda$ will be uniquely chosen to kill the average of the equation determining $\dot\varphi$ and the constant component of the given terms in the second and third equation that belong to the kernels of $A-I$ and $[A,\cdot\,]$ respectively, and solve the cohomological equations.

In the following we will repeatedly apply Lemmata 3.1–3.3 and Cauchy's inequality. Furthermore, we do not keep track of constants — just note that they may only depend on n and τ (from the Diophantine condition) and on $|G-\mathrm{id}|_{s+\sigma}$ and $|P-((\theta+\alpha),A\cdot r)|_{s+\sigma}$, and refer to them as C.

– First, the second equation has a solution

$$\dot R_0 = L^{-1}_{1,A}(\dot Q_0+\dot b-\bar b),$$

where $\bar b=\prod_{\mathrm{Ker}\,(A-I)}\int_{\mathbb{T}^n}\dot Q_0+\dot b/(2\pi)^n\,d\theta$, and

$$|\dot R_0|_s\le \frac{C}{\gamma^2\sigma^{\tau+n}}|\dot Q_0+\dot b|_{s+\sigma}.$$

– Second, we have

$$\dot\varphi(\theta+\alpha)-\dot\varphi(\theta)+p_1\cdot \dot R_0=\dot q_0+\dot\beta-\bar\beta,$$

where $\bar\beta=\int_{\mathbb{T}^n}\dot q_0-p_1\cdot R_0+\dot\beta/(2\pi)^n\,d\theta$; hence

$$\dot\varphi=L^{-1}_{\alpha}(\dot q_0+\dot\beta-\bar\beta),$$

satisfying

$$|\dot\varphi|_{s-\sigma}\le\frac{C}{\gamma^3\sigma^{2(\tau+n)}}|\dot q_0+\dot\beta|_{s+\sigma}.$$

– Third, the solution of the equation in $\dot R_1$ is

$$\dot R_1=L^{-1}_{2,A}(\widetilde Q_1+\dot B-\bar B),$$

where $\widetilde Q_1=\dot Q_1-(2P_2\cdot\dot R_0+\dot R_0'(\theta+\alpha)\cdot p_1)$, and $\bar B=\prod_{\mathrm{Ker}[A,\cdot\,]}\int_{\mathbb{T}^n}\widetilde Q_1+\dot B/(2\pi)^n\,d\theta$. It satisfies

$$|\dot R_1|_{s-2\sigma}\le\frac{C}{\gamma^2\sigma^{n+\tau}}|\widetilde Q_1+\dot B|_{s+\sigma}.$$

We now handle the unique choice of the correction $\delta\lambda=(\delta\beta,\delta b+\delta B\cdot r)$ given by $T_{\delta\lambda}$. Letting $\bar\lambda=(\bar\beta,\bar b+\bar B\cdot r)$, the map $f:\Lambda\to\Lambda,\ \delta\lambda\mapsto-\bar\lambda$, is well defined in the neighborhood of $\delta\lambda=0$.

In particular $f' = -\,\mathrm{id}$ when $G = \mathrm{id}$, and it will remain bounded away from 0 if G stays sufficiently close to the identity. In particular, $\delta\lambda \mapsto -\bar{\lambda}$ is affine: the system in Λ to solve $\lambda = 0$ is linear of the form $\int_{\mathbb{T}^n} a(G, \dot{Q}) + A(G)\cdot\delta\lambda = 0$, with diagonal close to 1 when G is close to the identity; hence f is invertible. Thus, there exists a unique $\delta\lambda$ such that $f(\delta\lambda) = 0$, satisfying

$$|\delta\lambda| \leq \frac{C}{\gamma^2\sigma^{\tau+n+1}}|\delta Q|_{G,s+\sigma}.$$

We finally have

$$|\dot{G}|_{s-2\sigma} \leq \frac{C}{\gamma^3}\,\frac{1}{\sigma^{2(\tau+n)+1}}|\delta Q|_{G,s+\sigma}.$$

Now, from the definition of $\dot{G} = G'^{-1}\cdot\delta G$, we get $\delta G = G'\cdot\dot{G}$. The unique solutions such that $\delta\varphi(0) = 0$, $\delta R_0(0) = 0$ and $\delta R_1(0) = 0$ are easily determined, since G is close to the identity and similar estimates hold for δG:

$$|\delta G|_{s-2\sigma} \leq \sigma^{-1}(1 + |G - \mathrm{id}|_s)\frac{C}{\sigma^{2(\tau+n)+1}}|\delta Q|_{G,s+\sigma}.$$

Finally, (4-3) uniquely determines δP.

Letting $\tau' = 2(\tau+n)+2$, up to redefining $\sigma' = \sigma/3$ and $s' = s+\sigma$, we have the stated estimates for all s', σ', where $s' < s' + \sigma'$. □

Proposition 4.3 (boundedness of ϕ''). *The bilinear map* $\phi''(x)$,

$$\phi''(x) : (T_G\mathcal{G}^{\sigma/n}_{s+\sigma}\times\vec{U}_{s+\sigma}\times\Lambda)^{\otimes 2} \to \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{\mathbb{C}}),$$

satisfies the estimates

$$|\phi''(x)\cdot\delta x^{\otimes 2}|_{G,s} \leq \frac{C''}{\sigma^{\tau''}}|\delta x|^2_{s+\sigma},$$

where $\tau'' \geq 1$ *and* C'' *is a constant depending on* $|x|_{s+\sigma}$.

Proof. Differentiating $\phi(x)$ twice yields

$$\begin{aligned}
-M\Big\{\big[&\delta G'\circ P\cdot\delta P + \delta G'\circ P\cdot\delta P + G''\circ P\cdot\delta P^2 - (\delta G'\circ P + G''\circ P\cdot\delta P)\cdot P'\cdot G'^{-1}\cdot\delta G\\
&\qquad\qquad\qquad\qquad - G'\circ P\cdot\big(\delta P'\cdot(-G'^{-1}\cdot\delta G'\cdot G'^{-1})\cdot\delta G\big)\big]\circ G^{-1}\\
&+\big[\delta G'\circ P\cdot\delta P + \delta G'\circ P\cdot\delta P + G''\circ P\cdot\delta P^2\\
&\qquad -(\delta G'\circ P + G''\circ P\cdot\delta P)\cdot P'\cdot G'^{-1}\cdot\delta G\\
&\qquad\qquad - G'\circ P\cdot\big(\delta P'\cdot(-G'^{-1}\cdot\delta G'\cdot G'^{-1})\cdot\delta G\big)\big]'\circ G^{-1}\cdot(-G'^{-1}\cdot\delta G)\circ G^{-1}\Big\}.
\end{aligned}$$

Once we precompose with G, the estimate follows. □

The hypotheses of Theorem A.1 are satisfied; hence the existence of (G, P, λ) with $Q = T_\lambda\circ G\circ P\circ G^{-1}$ is proved. Uniqueness and smoothness of the normal form follows from Propositions A.2–A.4. Theorem 4.1 follows, and hence Theorem A.

5. A generalization of Rüssmann's theorem

Theorem A provides a normal form that does not rely on any nondegeneracy assumption; thus, the existence of a translated Diophantine, reducible torus will be subordinated to eliminating the "parameters in excess" (β, B) using a nondegeneracy hypothesis. We will implicitly solve $B = 0$ and $\beta = 0$ by using the normal frequencies as *free* parameters and a torsion hypothesis respectively. Rüssmann's classical result will be the immediate small-dimensional case.

Elimination of B. Let $\Delta_m^s(\mathbb{R}) \subset \mathrm{GL}_m(\mathbb{R})$ be the open set of invertible matrices with simple, real eigenvalues. On $\mathbb{T}^n \times \mathbb{R}^m$, let us define

$$\widehat{U} = \bigcup_{A \in \Delta_m^s(\mathbb{R})} U(\alpha, A).$$

We recall that those $P \in U(\alpha, A)$ are diffeomorphisms of the form

$$P(\theta, r) = (\theta + \alpha + O(r), A \cdot r + O(r^2)),$$

on a neighborhood of $\mathbb{T}^n \times \{0\}$.

The following theorem is an intermediate, yet fundamental result to prove the translated torus theorem, Theorem C, and holds without requiring any torsion assumption on the class of diffeomorphisms.

Theorem 5.1 (twisted torus of codimension 1). *For every $P^0 \in U_{s+\sigma}(\alpha, A^0)$ with α Diophantine and $A^0 \in \Delta_m^s(\mathbb{R})$, there is a germ of C^∞-maps*

$$\psi : V_{s+\sigma} \to \mathcal{G}_s \times \widehat{U}_s \times \Lambda(\beta, b), \quad Q \mapsto (G, P, \lambda),$$

at $P^0 \mapsto (\mathrm{id}, P^0, 0)$ such that $Q = T_\lambda \circ G \circ P \circ G^{-1}$, where $\lambda = (\beta, b) \in \mathbb{R}^{n+1}$.

Corollary 5.2 (twisted torus). *If* 1 *does not belong to the spectrum of A^0, the translation correction b is* 0.

Proof. Denote by ϕ_A the operator ϕ, as now we want A to vary. Let us define the map

$$\hat{\psi} : \Delta_m^s(\mathbb{R}) \times V_{s+\sigma} \to \mathcal{G}_s \times \widehat{U}_s \times \Lambda, \quad (A, Q) \mapsto \hat{\psi}_A(Q) := \phi_A^{-1}(Q) = (G, P, \lambda),$$

in the neighborhood of (A^0, P^0) such that $Q = T_\lambda \circ G \circ P \circ G^{-1}$, where $\lambda = (\beta, b, B \cdot r)$, $\beta \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, such that $(A - I) \cdot b = 0$ and $B \in \mathrm{Mat}_m(\mathbb{R})$ satisfies $[B, A] = 0$. Equivalently, B is simultaneously diagonalizable with A, since A has simple spectrum; we can thus restrict our analysis to a neighborhood of A^0 in the subspace of those matrices commuting with A^0. Note that we can choose such a neighborhood so that it is contained in $\Delta_m^s(\mathbb{R})$. Then we study the dependence of B on A in their diagonal forms.

Without loss of generality, let A^0 be in its canonical form, and let Δ_{A^0} be the subspace of diagonal matrices, namely the matrices which commute with A^0. Consider the restriction of $\hat{\psi}$ to Δ_{A^0}. Let $A \in \Delta_{A^0}$ be close to A^0, let $\delta A := A^0 - A$ and write P^0 as

$$P^0(\theta, r) = \big(\theta + \alpha + O(r), (A^0 - \delta A) \cdot r + \delta A \cdot r + O(r^2)\big);$$

we remark that $P^0 = T_\lambda \circ P_A$, where

$$\lambda = \big(0, B(A) = (A^0 - A) \cdot A^{-1}\big), \quad [B(A), A] = 0,$$

and $P_A = (\theta + \alpha + O(r),\ A \cdot r + O(r^2))$, $A = A^0 - \delta A$.[2] Note that, since $A \in \Delta_{A^0}$ has simple spectrum, B is indeed in Δ_{A^0}.

According to Theorem A, $\phi_A(\mathrm{id}, P_A, \lambda) = P^0$; thus locally for all $A \in \Delta_{A^0}$ close to A^0 we have

$$\hat{\psi}(A, P^0) = (\mathrm{id}, P_A, B \cdot r), \quad B(A, P^0) = (A^0 - A) \cdot A^{-1} = \delta A \cdot (A^0 - \delta A)^{-1},$$

and, in particular $B(A^0, P^0) = 0$ and

$$\left.\frac{\partial B}{\partial A}\right|_{A=A^0} = -(A^0)^{-1},$$

which is invertible, since A^0 is so. Hence $A \mapsto B(A)$ is a local diffeomorphism on Δ_{A^0} and by the implicit function theorem (in finite dimension) locally for all Q close to P^0 there exists a unique $\bar{A}$ such that $B(\bar{A}, Q) = 0$. It remains to define $\psi(Q) = \hat{\psi}(\bar{A}, Q)$. □

The proof of Corollary 5.2 is immediate, by conditions (1-4).

Remark 5.3. This twisted-torus theorem relies on the peculiarity of the normal dynamics of the torus T_0^n. The direct applicability of the implicit function theorem is subordinated to the fact that no arithmetic condition is required on the characteristic (normal) frequencies so that the correction $A^0 + \delta A$ is well defined; beyond that, since having simple, real eigenvalues is an open property, the needed counter term B is indeed a diagonal matrix, so that the number of free frequencies (parameters) is enough to solve, implicitly, $B(A) = 0$. The generic case of complex eigenvalues is more delicate since one should guarantee that corrections $A^0 + \delta A$ at each step satisfy the Diophantine condition (1-2). It seems reasonable to think that one would need more parameters to control this issue, using the Whitney smoothness of ϕ on A, and verify that the measure of such stay positive; see [Féjoz 2004].

Elimination of β. If Q satisfies a torsion hypothesis, the existence of a translated Diophantine torus can be proved.

Theorem 5.4 (translated Diophantine torus). *Let α be Diophantine. On a neighborhood of $\mathbb{T}^n \times \{0\} \subset \mathbb{T}^n \times \mathbb{R}^n$, let $P^0 \in U(\alpha, A^0)$ be a diffeomorphism of the form*

$$P^0(\theta, r) = \big(\theta + \alpha + p_1(\theta) \cdot r + O(r^2),\ A^0 \cdot r + O(r^2)\big),$$

where A^0 is invertible and has simple, real eigenvalues and such that

$$\det\left(\int_{\mathbb{T}^n} p_1(\theta)\, d\theta\right) \neq 0.$$

If Q is close enough to P^0, there exists a unique A', close to A^0, and a unique $(G, P, b) \in \mathcal{G} \times U(\alpha, A') \times \mathbb{R}^n$ such that $Q = T_b \circ G \circ P \circ G^{-1}$.

Phrasing the thesis, the graph of $\gamma = R_0 \circ \varphi^{-1}$ is a translated torus on which the dynamics is conjugated to R_α by φ (remember the form of $G \in \mathcal{G}$ given in (1-3)). Before proceeding with the proof of Theorem 5.4, let us consider a parameter $c \in B_1^n(0)$ (the unit ball in $\mathbb{R}^n$) and the family of maps defined by $Q_c(\theta, r) :=$

[2] The terms $O(r^2)$ contain a factor $(I + \delta A \cdot A^{-1})^{-1}$.

$Q(\theta, c+r)$ obtained by translating the action coordinates. Considering the corresponding normal form operators ϕ_c, the parametrized version of Theorem A follows readily.

Now, if Q_c is close enough to P^0_c, Theorem 5.1 asserts the existence of $(G_c, P_c, \lambda_c) \in \mathcal{G} \times U(\alpha, A) \times \Lambda(\beta, b)$ such that

$$Q_c = T_\lambda \circ G_c \circ P_c \circ G_c^{-1}.$$

Hence we have a family of tori parametrized by $\tilde{c} = c + \int_{\mathbb{T}^n} \gamma/(2\pi)^n \, d\theta$,

$$Q(\theta, \tilde{c} + \tilde{\gamma}(\theta)) = \big(\beta(c) + \varphi \circ R_\alpha \circ \varphi^{-1}(\theta), b(c) + \tilde{c} + \tilde{\gamma}(\varphi \circ R_\alpha \circ \varphi^{-1}(\theta))\big),$$

where $\gamma := R_0 \circ \varphi^{-1}$ and $\tilde{\gamma} = \gamma - \int_{\mathbb{T}} \gamma/(2\pi) \, d\theta$.

Proof. Let $\hat{\varphi}$ be the function defined on $\mathbb{T}^n$ taking values in $\mathrm{Mat}_n(\mathbb{R})$ that solves the (matrix of) difference equation

$$\hat{\varphi}(\theta + \alpha) - \hat{\varphi}(\theta) + p_1(\theta) = \int_{\mathbb{T}^n} p_1(\theta) \frac{d\theta}{(2\pi)^n},$$

and let $F : (\theta, r) \mapsto (\theta + \hat{\varphi}(\theta) \cdot r, r)$. The diffeomorphism F restricts to the identity at T^n_0. At the expense of substituting P^0 and Q with $F \circ P^0 \circ F^{-1}$ and $F \circ Q \circ F^{-1}$ respectively, we can assume that

$$P^0(\theta, r) = \big(\theta + \alpha + p_1 \cdot r + O(r^2), A^0 \cdot r + O(r^2)\big), \quad p_1 = \int_{\mathbb{T}^n} p_1(\theta) \frac{d\theta}{(2\pi)^n}.$$

The germs so obtained from the initial P^0 and Q are close to one another.

The proof will follow from Theorem 5.1 and the elimination of the parameter $\beta \in \mathbb{R}^n$ obstructing the rotation conjugacy.

In line with the previous reasoning, we want to show that the map $c \mapsto \beta(c)$ is a local diffeomorphism. It suffices to show this for the trivial perturbation P^0_c. The Taylor expansion of P^0_c directly gives the normal form. In particular $b(c) = A^0 \cdot c + O(c^2)$, while the map $c \mapsto \beta(c) = p_1 \cdot c + O(c^2)$ is such that $\beta(0) = 0$ and $\beta'(0) = p_1$, which is invertible by twist hypothesis, and thus a local diffeomorphism. Hence, the analogous map for Q_c, which is a small C^1-perturbation, is a local diffeomorphism too and, together with Theorem 5.1, there exists unique $c \in \mathbb{R}^n$ and $A \in \mathrm{Mat}_n(\mathbb{R})$ such that $(\beta, B) = (0, 0)$. ☐

Remark 5.5. The theorem holds also on $\mathbb{T}^n \times \mathbb{R}^m$, with $m \geq n$, requiring that

$$\operatorname{rank}\left(\int_{\mathbb{T}^n} p_1(\theta) \, d\theta\right) = n.$$

This guarantees that $c \mapsto \beta(c)$ is submersive, but c solving $\beta(c) = 0$ would no more be uniquely determined.

Remark 5.6. Theorem 5.4 generalizes the classical translated curve theorem of Rüssmann in higher dimension, in the case of normally hyperbolic systems such that A has simple, real, nonzero eigenvalues, for general perturbations.

We stress the fact that if P^0 was of the form

$$P^0(\theta, r) = (\theta + \alpha + O(r), I \cdot r + O(r^2)),$$

like in the original frame studied by Rüssmann, we would need a whole matrix $B \in \mathrm{Mat}_n(\mathbb{R})$ in order to solve the homological equations, and, having just n characteristic frequencies at our disposal, it is hopeless to completely solve $B = 0$ and eliminate the whole obstruction. The torus would not be just translated.

Appendix A: The inverse function theorem and regularity of ϕ

We state here the implicit function theorem we use to prove Theorem A as well as the regularity statements needed to guarantee uniqueness and smoothness of the normal form. These results follow from [Féjoz 2010; 2017]. Note that we endowed functional spaces with weighted norms, and bounds appearing in Propositions 4.2–4.3 may depend on $|x|_s$ (as opposed to the analogous statements in [Féjoz 2010; 2017]); for the corresponding proofs taking into account these (slight) differences, we send the reader to [Massetti 2015a; 2015b] and the proof or Moser's theorem therein.

Let $E = (E_s)_{0<s<1}$ and $F = (F_s)_{0<s<1}$ be two decreasing families of Banach spaces with increasing norms $|\cdot|_s$ and let $B_s^E(\sigma) = \{x \in E : |x|_s < \sigma\}$ be the ball of radius σ centered at 0 in E_s.

On account of composition operators, we additionally endow F with some deformed norms which depend on $x \in B_s^E(s)$ such that

$$|y|_{0,s} = |y|_s \quad \text{and} \quad |y|_{\hat{x},s} \le |y|_{x,s+|x-\hat{x}|_s}.$$

Consider then operators commuting with inclusions $\phi : B_{s+\sigma}^E(\sigma) \to F_s$, with $0 < s < s+\sigma < 1$, such that $\phi(0) = 0$.

We then suppose that if $x \in B_{s+\sigma}^E(\sigma)$ then $\phi'(x) : E_{s+\sigma} \to F_s$ has a right inverse $\phi'^{-1}(x) : F_{s+\sigma} \to E_s$ (for the particular operators ϕ of this work, ϕ' is both left- and right-invertible).

Suppose ϕ is at least twice differentiable.

Let $\tau := \tau' + \tau''$ and $C := C'C''$.

Theorem A.1 (inverse function theorem). *Assume*

$$|\phi'^{-1}(x) \cdot \delta y|_s \le \frac{C'}{\sigma^{\tau'}} |\delta y|_{x,s+\sigma}, \tag{A-1}$$

$$|\phi''(x) \cdot \delta x^{\otimes 2}|_{x,s} \le \frac{C''}{\sigma^{\tau''}} |\delta x|_{s+\sigma}^2 \quad \forall s, \sigma : 0 < s < s+\sigma < 1, \tag{A-2}$$

where C' and C'' depend on $|x|_{s+\sigma}$, and $\tau', \tau'' \ge 1$.

For any s, σ, η with $\eta < s$ and $\varepsilon \le \eta\sigma^{2\tau}/(2^{8\tau}C^2)$ ($C \ge 1$, $\sigma < 3C$), ϕ has a right inverse $\psi : B_{s+\sigma}^F(\varepsilon) \to B_s^E(\eta)$. In other words, ϕ is locally surjective:

$$B_{s+\sigma}^F(\varepsilon) \subset \phi(B_s^E(\eta)).$$

Proposition A.2 (Lipschitz continuity of ψ). *Let $\sigma < s$. If $y, \hat{y} \in B_{s+\sigma}^F(\varepsilon)$ with $\varepsilon = 3^{-4\tau}2^{-16\tau}\sigma^{6\tau}/(4C^3)$, the following inequality holds*:

$$|\psi(y) - \psi(\hat{y})|_s \le L|y - \hat{y}|_{x,s+\sigma},$$

with $L = 2C'/\sigma^{\tau'}$. In particular, ψ being the unique local right inverse of ϕ, it is also its unique left inverse.

Proposition A.3 (smooth differentiation of ψ). *Let $\sigma < s < s+\sigma$ and ε be as in Proposition A.2. There exists a constant K such that for every $y, \hat{y} \in B^F_{s+\sigma}(\varepsilon)$ we have*

$$|\psi(\hat{y}) - \psi(y) - \phi'^{-1}(\psi(y))(\hat{y}-y)|_s \le K(\sigma)|\hat{y}-y|^2_{x,s+\sigma},$$

and the map $\psi' : B^F_{s+\sigma}(\varepsilon) \to L(F_{s+\sigma}, E_s)$ defined locally by $\psi'(y) = \phi'^{-1}(\psi(y))$ is continuous. In particular ψ has the same degree of smoothness as ϕ.

It is sometimes convenient to extend ψ to non-Diophantine characteristic frequencies (α, A). Whitney smoothness guarantees that such an extension exists. Let suppose that $\phi(x) = \phi_\nu(x)$ depends on some parameter $\nu \in B^k$ (the unit ball of $\mathbb{R}^k$) and that it is C^1 with respect to ν and that estimates on $\phi_\nu'^{-1}$ and ϕ_ν'' are uniform with respect to ν over some closed subset D of $\mathbb{R}^k$.

Proposition A.4 (Whitney differentiability). *Let us fix ε, σ, s as in Proposition A.2. The map $\psi : D \times B^F_{s+\sigma}(\varepsilon) \to B^E_s(\eta)$ is C^1-Whitney differentiable and extends to a map $\psi : \mathbb{R}^{2n} \times B^F_{s+\sigma}(\varepsilon) \to B^E_s(\eta)$ of class C^1. If ϕ is C^k, $1 \le k \le \infty$, with respect to ν, this extension is C^k.*

Appendix B: Inversion of a holomorphism of $\mathbb{T}^n_s$

We present here a classical result and a lemma that justify the well-definedness of the normal form operator ϕ defined in Section 1.

Complex extensions of manifolds are defined with the help of the ℓ^∞-norm.

Let

$$\mathbb{T}^n_{\mathbb{C}} = \mathbb{C}^n/2\pi\mathbb{Z}^n \quad \text{and} \quad \mathrm{T}^n_{\mathbb{C}} = \mathbb{T}^n_{\mathbb{C}} \times \mathbb{C}^m,$$
$$\mathbb{T}^n_s = \big\{\theta \in \mathbb{T}^n_{\mathbb{C}} : |\theta| := \max_{1\le j\le n} |\operatorname{Im}\theta_j| \le s\big\}, \quad \mathrm{T}^n_s = \big\{(\theta, r) \in \mathrm{T}^n_{\mathbb{C}} : |(\operatorname{Im}\theta, r)| \le s\big\},$$

where $|(\operatorname{Im}\theta, r)| := \max_{1\le j\le n} \max(|\operatorname{Im}\theta_j|, |r_j|)$.

Let also define $\mathbb{R}^n_s := \mathbb{R}^n \times (-s, s)$ and consider the universal covering of $\mathbb{T}^n_s$, $p : \mathbb{R}^n_s \to \mathbb{T}^n_s$.

Theorem B.1. *Let $v : \mathbb{T}^n_s \to \mathbb{C}^n$ be a vector field such that $|v|_s < \sigma/n$. The map $\mathrm{id} + v : \mathbb{T}^n_{s-\sigma} \to \mathbb{R}^n_s$ induces a map $\varphi = \mathrm{id} + v : \mathbb{T}^n_{s-\sigma} \to \mathbb{T}^n_s$ which is a biholomorphism and there is a unique biholomorphism $\psi : \mathbb{T}^n_{s-2\sigma} \to \mathbb{T}^n_{s-\sigma}$ such that $\varphi \circ \psi = \mathrm{id}_{\mathbb{T}^n_{s-2\sigma}}$.*

In particular the following hold:

$$|\psi - \mathrm{id}|_{s-2\sigma} \le |v|_{s-\sigma}$$

and, if $|v|_s < \sigma/(2n)$,

$$|\psi' - \mathrm{id}|_{s-2\sigma} \le \frac{2}{\sigma}|v|_s.$$

For the proof we again direct readers to [Massetti 2015a; 2015b].

Corollary B.2 (well-definedness of the normal form operator ϕ). *For all s, σ if $G \in \mathcal{G}^{\sigma/n}_{s+\sigma}$, then $G^{-1} \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{s+\sigma})$.*

Proof. We recall the form of $G \in \mathcal{G}^{\sigma/n}_{s+\sigma}$:

$$G(\theta, r) = (\varphi(\theta), R_0(\theta) + R_1(\theta)\cdot r).$$

G^{-1} reads as

$$G^{-1}(\theta, r) = \big(\varphi^{-1}(\theta),\, R_1^{-1} \circ \varphi^{-1}(\theta) \cdot (r - R_0 \circ \varphi^{-1}(\theta))\big).$$

Up to rescaling norms by a factor of $\frac{1}{2}$, like $\|x\|_s := \frac{1}{2}|x|$, the statement is straightforward and follows from Theorem B.1. By abuse of notations, we keep on denoting $\|x\|_s$ by $|x|_s$. □

Appendix C: Fourier norms

Let $\mathcal{A}(\mathrm{T}^n_s, \mathbb{C})$ be the space of holomorphic functions on T^n_s with values in $\mathbb{C}$, endowed with the norm

$$\|f\|_s = \sum_k \sup_{|r|<s} |f_k(r)|\, e^{|k|s}, \quad |k| = |k_1| + \cdots + |k_n|.$$

If $f \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$, the definition of the norm is adapted in the obvious way and the expression $|f_k(r)|$ denotes the standard operator norm $\sup_{|\xi|=1} |f_k(r)\xi|$. If $f : \mathrm{T}^n_s \to \mathbb{C}^n$, then $\|f\|_s = \max_{1\le j\le n}(\|f^j\|_s)$.

Lemma C.1. *Let $f \in \mathcal{A}(\mathrm{T}^n_{s+\sigma}, \mathbb{C})$ and let $h \in \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^n)$ be such that $\|h\|_s < \sigma/e$, then*

$$\|f(\theta, r + h(\theta, r))\|_s \le \frac{1}{1 - e\|h\|_s/\sigma} \|f\|_{s+\sigma}.$$

Proof. Let $f(\theta, r + h(\theta, r)) = \sum_n D^n f(\theta, r) h^n(\theta, r)/n!$ be the Taylor expansion of f. Then

$$\|f(\theta, r + h(\theta, r))\|_s \le \sum_k \sup_{|r|<s} \Bigg(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} |D^n f_\ell(r)|\, |h_{k^1}(r)| \cdots |h_{k^n}(r)| \Bigg) e^{|k|s},$$

where $k^i \in \mathbb{Z}^n$, $i = 1, \dots, n$, are the Fourier indexes. Since $|k| \le |\ell| + |k^1| + \cdots + |k^n|$,

$$\begin{aligned}
\|f(\theta, r + h(\theta, r))\|_s &\le \sum_k \sup_{|r|<s} \Bigg(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} |D^n f_\ell(r)| e^{|\ell|s}\, |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \Bigg) \\
&\le \sum_k \Bigg(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s}\, |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \Bigg) \\
&\le \sum_k \Bigg(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s} |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \Bigg) \\
&\le \sum_n \frac{1}{n!} \sum_\ell \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s} \sum_{k^1} \sup_{|r|<s} |h_{k^1}(r)| e^{|k^1|s} \cdots \sum_{k^n} \sup_{|r|<s} |h_{k^n}(r)| e^{|k^n|s} \\
&\le \sum_\ell \Bigg(\sum_n \frac{n^n}{n!} \sup_{|r|<s+\sigma} |f_\ell(r)| \frac{\|h\|_s^n}{\sigma^n} \Bigg) e^{|\ell|(s+\sigma)},
\end{aligned}$$

where the last estimate follows from the fact that $(D^n f)_\ell = D^n(f_\ell)$ and the classical Cauchy's estimate by observing that for all $|r| < s$ letting $\mathbb{R}^n_{s+\sigma} \ni \xi \ne 0$, the analytic function $\varphi(t) = f(r + t\xi)$ on the complex disc $|t| < \sigma/|\xi|$ satisfies $d^n\varphi/dt^n|_{t=0} = D^n f(r)\xi^n$. The factor n^n comes from the classical bound on the norm of a symmetric multilinear mapping by the associated homogeneous polynomial; see for example [Harris 1975].

It thus follows that

$$\begin{aligned}\|f(\theta, r+h(\theta, r))\|_s &\le 2\sum_{\ell} \sup_{|r|<s+\sigma} |f_\ell(r)|_{s+\sigma} \sum_{n\ge 1} \frac{e^n}{\sqrt{2\pi n}} \left(\frac{\|h\|_s}{\sigma}\right)^n e^{|\ell|(s+\sigma)} \\ &\le 2\|f\|_{s+\sigma} \frac{1}{2(1-e\|h\|_s/\sigma)};\end{aligned}$$

hence the stated bound. □

Lemma C.2. *Let $f \in \mathcal{A}(\mathrm{T}^n_{s+\sigma}, \mathbb{C})$ and $h \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C})$ be such that $\|h\|_s < \sigma$; then*

$$\|f(\theta + h(\theta), r)\|_s \le \|f\|_{s+\sigma}.$$

For the proof, see [Chierchia 2003, Appendix B] for example.

Acknowledgments

The author has been supported by the ERC starting grant HamPDEs. The author is grateful for the mathematical (and moral) support and advice of A. Chenciner and J. Féjoz all the way through her Ph.D, at the Astronomie et Systèmes Dynamiques group of the IMCCE, Observatoire de Paris. This work owes a lot to the meaningful talks of those times. Thank you also to L. Biasco, T. Castan, B. Fayad, J. Laskar, J.-P. Marco, P. Mastrolia, L. Niederman, P. Robutel, A. Pousse, and M. Procesi for the enlightening discussions we had and their constant interest in this (and future) work.

References

[Arnol'd 1961] V. I. Arnol'd, "Small denominators, I: Mapping the circle onto itself", *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86. In Russian. MR Zbl

[Bost 1986] J.-B. Bost, "Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnol'd, Moser, Rüssmann, Zehnder, Herman, Pöschel, . . .)", exposé 639 , pp. 113–157 in *Seminar Bourbaki,* 1984/85, Astérisque **133-134**, Société Mathématique de France, Paris, 1986. MR Zbl

[Calleja et al. 2013] R. C. Calleja, A. Celletti, and R. de la Llave, "A KAM theory for conformally symplectic systems: efficient algorithms and their validation", *J. Differential Equations* **255**:5 (2013), 978–1049. MR Zbl

[Chenciner 1985a] A. Chenciner, "Bifurcations de points fixes elliptiques, I: Courbes invariantes", *Inst. Hautes Études Sci. Publ. Math.* **61** (1985), 67–127. MR Zbl

[Chenciner 1985b] A. Chenciner, "Bifurcations de points fixes elliptiques, II: Orbites periodiques et ensembles de Cantor invariants", *Invent. Math.* **80**:1 (1985), 81–106. MR Zbl

[Chenciner 1988] A. Chenciner, "Bifurcations de points fixes elliptiques, III: Orbites périodiques de "petites" périodes et élimination résonnante des couples de courbes invariantes", *Inst. Hautes Études Sci. Publ. Math.* **66** (1988), 5–91. MR Zbl

[Chierchia 2003] L. Chierchia, "KAM lectures", pp. 1–55 in *Dynamical systems, I*, Scuola Norm. Sup., Pisa, 2003. MR Zbl

[Eliasson et al. 2015] L. H. Eliasson, B. Fayad, and R. Krikorian, "Around the stability of KAM tori", *Duke Math. J.* **164**:9 (2015), 1733–1775. MR Zbl

[Féjoz 2004] J. Féjoz, "Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman)", *Ergodic Theory Dynam. Systems* **24**:5 (2004), 1521–1582. MR Zbl

[Féjoz 2010] J. Féjoz, *Mouvements périodiques et quasi-périodiques dans le problème des n corps*, Mémoire d'habilitation, Université Pierre et Marie Curie, 2010, available at https://tel.archives-ouvertes.fr/tel-00702650.

[Féjoz 2017] J. Féjoz, "Introduction to KAM theory, with a view to celestial mechanics", pp. 387–433 in *Variational methods in imaging and geometric control*, edited by M. Bergounioux et al., Radon Series on Computational and Applied Mathematics **18**, de Gruyter, Berlin, 2017.

[Harris 1975] L. A. Harris, "Bounds on the derivatives of holomorphic functions of vectors", pp. 145–163 in *Analyse fonctionnelle et applications* (Rio de Janeiro, 1972), Actualités Aci. Indust. **1367**, Hermann, Paris, 1975. MR Zbl

[Herman 1983] M.-R. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau, Vol. 1*, Astérisque **103**, Société Mathématique de France, Paris, 1983. MR Zbl

[Herman and Sergeraert 1971] M. Herman and F. Sergeraert, "Sur un théorème d'Arnold et Kolmogorov", *C. R. Acad. Sci. Paris Sér. A-B* **273** (1971), A409–A411. MR Zbl

[Kolmogorov 1954] A. N. Kolmogorov, "On conservation of conditionally periodic motions for a small change in Hamilton's function", *Dokl. Akad. Nauk SSSR* (*N.S.*) **98** (1954), 527–530. In Russian. MR Zbl

[Massetti 2015a] J. E. Massetti, "Normal forms for perturbations of systems possessing a diophantine invariant torus", preprint, 2015. arXiv

[Massetti 2015b] J. E. Massetti, *Quasi-périodicité et quasi-conservativité*, Ph.D. thesis, Observatoire de Paris, 2015.

[Meyer 1975] K. R. Meyer, "The implicit function theorem and analytic differential equations", pp. 191–208 in *Dynamical systems* (Warwick, 1974), edited by A. Manning, Lecture Notes in Mathematics **468**, Springer, 1975. MR Zbl

[Moser 1962] J. Moser, "On invariant curves of area-preserving mappings of an annulus", *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1962** (1962), 1–20. MR Zbl

[Moser 1966] J. Moser, "A rapidly convergent iteration method and non-linear differential equations, II", *Ann. Scuola Norm. Sup. Pisa* (3) **20** (1966), 499–535. MR Zbl

[Moser 1967] J. Moser, "Convergent series expansions for quasi-periodic motions", *Math. Ann.* **169** (1967), 136–176. MR Zbl

[Rüssmann 1970] H. Rüssmann, "Kleine Nenner, I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes", *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1970** (1970), 67–105. MR Zbl

[Rüssmann 1976] H. Rüssmann, "On optimal estimates for the solutions of linear difference equations on the circle", *Celestial Mech.* **14**:1 (1976), 33–37. MR Zbl

[Sevryuk 1999] M. B. Sevryuk, "The lack-of-parameters problem in the KAM theory revisited", pp. 568–572 in *Hamiltonian systems with three or more degrees of freedom* (S'Agaró, 1995), edited by C. Simó, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **533**, Kluwer Acad. Publ., Dordrecht, 1999. MR Zbl

[Wagener 2010] F. Wagener, "A parametrised version of Moser's modifying terms theorem", *Discrete Contin. Dyn. Syst. Ser. S* **3**:4 (2010), 719–768. MR Zbl

[Yoccoz 1992] J.-C. Yoccoz, "Travaux de Herman sur les tores invariants", exposé 754, pp. 311–344 in *Séminaire Bourbaki,* 1991/92, Astérisque **206**, Société Mathématique de France, Paris, 1992. MR Zbl

[Zehnder 1975] E. Zehnder, "Generalized implicit function theorems with applications to some small divisor problems, I", *Comm. Pure Appl. Math.* **28** (1975), 91–140. MR Zbl

Received 31 Aug 2016. Revised 14 Jun 2017. Accepted 28 Jul 2017.

JESSICA ELISA MASSETTI: jmassetti@mat.uniroma3.it
Università degli Studi Roma Tre, Dipartimento di Matematica e Fisica, Sezione Matematica, Rome, Italy

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.171

GLOBAL RESULTS FOR EIKONAL HAMILTON–JACOBI EQUATIONS ON NETWORKS

ANTONIO SICONOLFI AND ALFONSO SORRENTINO

We study a one-parameter family of eikonal Hamilton–Jacobi equations on an embedded network, and prove that there exists a unique critical value for which the corresponding equation admits global solutions, in a suitable viscosity sense. Such a solution is identified, via a Hopf–Lax-type formula, once an admissible trace is assigned on an *intrinsic boundary*. The salient point of our method is to associate to the network an *abstract graph*, encoding all of the information on the complexity of the network, and to relate the differential equation to a *discrete functional equation* on the graph. Comparison principles and representation formulae are proven in the supercritical case as well.

1. Introduction

Over the last few years there has been an increasing interest in the study of the Hamilton–Jacobi equation on networks and related questions. These problems, in fact, involve a number of subtle theoretical issues and have a great impact in the applications in various fields, for example, to data transmission, traffic management problems, etc. While locally — i.e., on each branch of the network (*arcs*) — the study reduces to the analysis of 1-dimensional problems, the main difficulties arise in matching together the information "converging" at the *juncture* of two or more arcs, and relating the *local* analysis at a juncture with the *global* structure/topology of the network.

In this article, we provide a thorough discussion of the above issues in the case of eikonal-type Hamilton–Jacobi equations on embedded networks (in $\mathbb{R}^n$ or on a Riemannian manifold, see Remark 3.1). We show that there exists a unique critical value for which the corresponding equation admits global solutions, and extend most of the results known in the continuous setting for the critical and supercritical cases. More specifically: we determine a uniqueness set (the *Aubry set*) for global solutions and provide -1-Lax-type representation formulae; we study critical subsolutions, their properties and constraints, and show the existence of C^1 critical subsolutions; we describe -1-Lax representation formulae for maximal supercritical subsolutions. See the Main Theorem in Section 4 for a more detailed description.

The main rationale behind our approach consists in neatly distinguishing between the local problem on the arcs and the global analysis on the network. While the former can be solved by means of (classical) 1-dimensional viscosity techniques, the latter is definitely more engaging.

MSC2010: primary 35F21, 35R02; secondary 35B51, 49L25.
Keywords: Hamilton–Jacobi equation, embedded networks, graphs, viscosity solutions, viscosity subsolutions, comparison principle, discrete functional equation on graphs, Hopf–Lax formula, discrete weak KAM theory.

Our novel idea is to tackle it by associating to the network an *abstract graph*, encoding all of the information on the complexity of the network, and to relate the problem to a *discrete functional equation* on the graph. This allows us to pursue a global analysis of the equation — that goes beyond what happens at a single juncture — as well as to prove uniqueness and comparison principles in a simpler way. To the best of our knowledge, this is the first time that comparison-type results are obtained in the network setting by completely bypassing the difficulties involved in the Crandall–Lions doubling variable method, in favor of a more direct analysis of a discrete equation.

In addition to this, by exploiting the simple geometry of the abstract graph we are able to identify an intrinsic boundary — the *Aubry set* — on which admissible traces can be assigned in order to get unique critical solutions on the whole network; these solutions can be represented by means of Hopf–Lax-type formulae. In the supercritical case we get existence and uniqueness of solutions, on any open subset of the network, continuously extending admissible data prescribed on the complement.

Let us point out that the problem of formulating boundary problems on the network and accordingly determining "natural" subsets on which to assign boundary data is a subtle issue, yet not well settled in the literature; we believe that our approach helps clarify this matter, at least in the class of equations that we are considering.

The notions of viscosity solution and subsolution that we adopt are very natural in this setting (see Definitions 3.6 and 3.7). More specifically, the tests we use at vertices are classical in viscosity solutions theory and consist in (unilateral) state-constraint-type boundary conditions, introduced by Soner [1986] to study control problems with constraints. In this regard, the notion of solution requires that at each vertex the state-constraint condition holds for at least one arc ending there: it does not require other mixing conditions (on the vertices) between equations defined on different incident arcs.

Very recently, the same notion of solution has been also considered by Lions and Souganidis [2016] to deal with 1-dimensional junction-type problems for nonconvex discounted Hamilton–Jacobi equations and study its well-posedness (i.e., comparison principle and existence). Global solutions on networks, however, are not therein studied.

As far as subsolutions are concerned, we only ask that they are continuous on the network and are (viscosity) subsolutions to the equation on the interior of each arc; no extra conditions are required on vertices. These assumptions are the minimal requirements that one needs to ask and, at a first sight, it might seem surprising that they are sufficient to develop a significant global theory. However, the validity of this approach is supported, among other things, by the fact that the notion of solutions can be recovered in terms of a maximal subsolution attaining a specific value at a given point (vertex or internal point); see Theorem 7.1.

We also wish to point out that our hypotheses, both on the topology of network and the Hamiltonians, are very general. As far as the network is concerned, we only ask it to be made up by finite arcs and connected; hence, it may well include multiple connections between different vertices, as well as the presence of loops.

The Hamiltonians are assumed continuous in both variables, quasiconvex and coercive in the first-order variable on any arc. Hamiltonians on different arcs are independent from each other and no compatibility conditions at the vertices are required. See Section 3B for more details.

We are confident that this very same set of ideas can be successfully applied to a broad range of other problems, for example, to the study of the *discounted* Hamilton–Jacobi equation on networks or to prove *homogenization* results for the Hamilton–Jacobi equation on periodic networks (also known as *topological crystals*). We plan to address these and other questions in a future work (in preparation).

1A. ***Previous related literature.*** There is a huge amount of literature related to differential equations on networks, or other nonregular geometric structures (ramified/stratified spaces), in various contexts: hyperbolic problems, traffic flows, evolutionary equations, (regional) control problems, Hamilton–Jacobi equations, etc. An exhaustive description of all of these areas would go well beyond the aims of this paper; we mention a few noteworthy papers, [Achdou et al. 2013; Barles et al. 2013; 2014; Bressan and Hong 2007; Camilli and Marchi 2013; Camilli et al. 2013; Davini et al. 2016; Galise et al. 2015; Garavello and Piccoli 2006; Imbert and Monneau 2016; 2017; Imbert et al. 2013; Lions and Souganidis 2016; Pokornyi and Borovskikh 2004; Rao et al. 2014; Schieborn and Camilli 2013; Soner 1986].

A model similar to ours has been previously considered by Schieborn and Camilli [2013], however, just in the supercritical case and under some restriction on the topology of the network. In comparison with their hypothesis, we do not require continuity of the Hamiltonians at the vertices (and accordingly, no mixed conditions on the test functions at the vertices) and we do not ask a priori existence of a regular strict subsolution.

Other relevant recent contributions are [Lions and Souganidis 2016] (which we have already mentioned above) and [Imbert and Monneau 2017]. In particular, the latter is a substantial work — whose point of view and techniques are rather different from ours — in which Imbert and Monneau attempt to recover the doubling variable method to their setting, by introducing an extra parameter (the flux limiter) and a companion equation (the junction condition) and by using special vertex test functions. See also other related works by the same authors and collaborators [Galise et al. 2015; Imbert et al. 2013; Imbert and Monneau 2016].

Our analysis of the discrete functional equation is based on ideas and techniques inspired by the so-called *weak KAM theory*, first developed by Fathi [2008] for the study of Tonelli Hamiltonian systems on closed manifolds; see also [Sorrentino 2015]. Developing a similar approach in the discrete setting is very natural and has been already exploited in several other works. In [Bernard and Buffoni 2006; 2007], for example, a discretization of weak KAM theory was applied to investigate the properties of optimal transport maps; a more systematic development of a discrete weak KAM theory for *cost functions* was described by Zavidovique [2010; 2012]; see also [Davini et al. 2016]. In particular, [Zavidovique 2012] shares ideas similar to ours, although our setting has the peculiarity of this interplay between the discrete structure and the embedded network.

From a more dynamical systems point of view, a discrete analogue of Aubry–Mather theory and weak KAM theory was also discussed in [Gomes 2005]; see [Su and Thieullen 2015] for a recent related work.

1B. ***Organization of the article.*** The article is organized as follows.

In Section 2, we provide a brief introduction to some topics in graph theory that will be needed in the following.

In Section 3, we describe our setting and the main objects involved in our analysis. More specifically: in Section 3A we introduce the concept of embedded network and its properties; in Section 3B we define Hamiltonians on a network and we detail which hypotheses we will be imposing thereafter; see (Hγ1)–(Hγ4). Finally, in Section 3C we introduce the eikonal Hamilton–Jacobi equation on a network ($\mathcal{H}Ja$) and provide suitable notions for viscosity solutions and subsolutions (see Definitions 3.6 and 3.7).

Section 4 provides a statement of our main results (see the Main Theorem) and an outline of the strategy of the proof, in order to guide the reader through Section 5 (local part), Section 6 (global part) and Section 7 (from global to local part).

2. Preliminaries on graph theory

We recall some basic material on the theory of abstract graphs and on functions defined on them. For a more detailed presentation of these and other related topics, we refer the interested reader, for instance, to [Sunada 2013].

2A. ***Abstract graphs.*** A (abstract) graph $\boldsymbol{X} = (\boldsymbol{V}, \boldsymbol{E})$ is an ordered pair of disjoint sets $\boldsymbol{V}$ and $\boldsymbol{E}$, which are called, respectively, *vertices* and (directed) *edges*, plus two functions

$$\mathrm{o} : \boldsymbol{E} \to \boldsymbol{V}$$

and

$$\begin{aligned} {}^{-} : \boldsymbol{E} &\to \boldsymbol{E}, \\ e &\mapsto \bar{e}, \end{aligned}$$

with the latter assumed to be a fixed-point-free involution, namely satisfying

$$\bar{e} \neq e \quad \text{and} \quad \bar{\bar{e}} = e \quad \text{for any } e \in \boldsymbol{E}.$$

We give the following geometric picture of the setting: $\mathrm{o}(e)$ is the *origin* (initial vertex) of e and $\bar{e}$ is its *reversed* edge, namely the same edge but with the opposite orientation. Analogously we define

$$\mathrm{t}(e) = \mathrm{o}(\bar{e}),$$

the *terminal* vertex of e. The following compatibility condition holds true:

$$\mathrm{t}(\bar{e}) = \mathrm{o}(\bar{\bar{e}}) = \mathrm{o}(e).$$

We say that e links $\mathrm{o}(e)$ to $\mathrm{t}(e)$; observe that it might well happen that $\mathrm{o}(e) = \mathrm{t}(e)$, and in this case e will be called a *loop*. An edge is also said to be incident on $\mathrm{o}(e)$ and $\mathrm{t}(e)$. Two vertices are called *adjacent* if there is an edge linking them or, in other terms, if there is an edge incident on both of them.

We say that the graph is *finite* if the set $\boldsymbol{E}$, and consequently $\boldsymbol{V}$, has a finite number of elements. We denote by $|\boldsymbol{V}|$ and $|\boldsymbol{E}|$ the number of vertices and edges.

We define a *path* to be a finite sequence of concatenated edges, namely $\xi = (e_1, \dots, e_M) = (e_i)_{i=1}^{M}$ satisfying

$$\mathrm{t}(e_j) = \mathrm{o}(e_{j+1}) \quad \text{for any } j = 1, \dots, M-1.$$

We set $\mathrm{o}(\xi) = \mathrm{o}(e_1)$ and $\mathrm{t}(\xi) = \mathrm{t}(e_M)$ and call them the initial and final vertices of the path. We say that ξ links $\mathrm{o}(\xi)$ to $\mathrm{t}(\xi)$; we also say that ξ is incident on some vertex if there is some edge composing the path incident on it.

Given two paths ξ and η, we say that ξ is contained in η, mathematically $\xi \subset \eta$, if the edges of ξ make up a subset of the edges of η. If such a subset is proper, we say that ξ is *properly contained* in η. If $\mathrm{t}(\xi) = \mathrm{o}(\eta)$, we denote by $\xi \cup \eta$ the path obtained via concatenation of ξ and η.

We call a path a *loop* or a *cycle* if $\mathrm{o}(\xi) = \mathrm{t}(\xi)$. A path without repetition of vertices except possibly the initial and terminal ones will be called *simple*; in other terms $\xi = (e_i)_1^M$ is simple if

$$\mathrm{t}(e_i) = \mathrm{t}(e_j) \implies i = j,$$

or if there are no cycles properly contained in ξ. Note that there are finitely many simple paths in a finite graph.

A graph is called *connected* if any two vertices are linked by some path. All of the graphs we will consider hereafter are understood to be connected and finite. Observe that the connectedness assumption implies that the map o (and hence t) is surjective.

Given $x \in \boldsymbol{V}$, we set

$$\boldsymbol{E}_x = \{e \in \boldsymbol{E} \mid \mathrm{o}(e) = x\}, \tag{1}$$

which we call $\boldsymbol{E}_x$, the *star centered at* x; it should be considered as a sort of tangent space to the graph at x. The cardinality of $\boldsymbol{E}_x$ is called the *degree* (or *valence*) of the vertex x.

2B. ***Functions on graphs.*** In the following we will be interested in functions defined on abstract graphs. It is useful to introduce the following notions:

- We define the 0-*cochain group* $\mathfrak{C}^0(\boldsymbol{X}, \mathbb{R})$ as the space of functions from $\boldsymbol{V}$ to $\mathbb{R}$.
- We define the 1-*cochain group* $\mathfrak{C}^1(\boldsymbol{X}, \mathbb{R})$ as the space of functions from $\boldsymbol{E}$ to $\mathbb{R}$, with the compatibility condition $\omega(\bar{e}) = -\omega(e)$. This space plays the role of 1-forms on the graph. From now on we will indicate the reverse edge $\bar{e}$ by $-e$ and we will consider the pairing $\langle \omega, e \rangle := \omega(e)$.

The relation between $\mathfrak{C}^0(\boldsymbol{X}, \mathbb{R})$ and $\mathfrak{C}^1(\boldsymbol{X}, \mathbb{R})$ can be expressed in terms of the so-called *coboundary operator*, or *differential*, $\mathrm{d} : \mathfrak{C}^0(\boldsymbol{X}, \mathbb{R}) \to \mathfrak{C}^1(\boldsymbol{X}, \mathbb{R})$, which is defined for any $f \in \mathfrak{C}^0(\boldsymbol{X}, \mathbb{R})$ and $e \in \boldsymbol{E}$ as

$$\mathrm{d}f(e) := f(\mathrm{t}(e)) - f(\mathrm{o}(e)).$$

We can embed these spaces with the standard topology. A notion of convergence on the cochain spaces is given via

$$f_n \to f \iff f_n(x) \to f(x) \quad \text{for any } x \in \boldsymbol{V},$$
$$\omega_n \to \omega \iff \omega_n(e) \to \omega(e) \quad \text{for any } e \in \boldsymbol{E}.$$

A sequence f_n is said to be *equibounded* if

$$|f_n(x)| \leq \beta \quad \text{for any } x \in \boldsymbol{V} \text{ and some } \beta > 0;$$

similarly ω_n is said equibounded if

$$|\langle \omega_n, e\rangle| \le \beta \quad \text{for any } e \in \boldsymbol{E} \text{ and some } \beta > 0.$$

It is clear that any equibounded sequences f_n and ω_n are convergent, up to subsequences.

We directly deduce from the above definitions:

Proposition 2.1. *Let f_n and f be in $\mathfrak{C}^0(\boldsymbol{X}, \mathbb{R})$:*

(i) *If $f_n \to f$, then $\mathrm{d}f_n \to \mathrm{d}f$.*

(ii) *If $\mathrm{d}f_n$ is equibounded and the sequence $f_n(x_0)$ is bounded for some vertex x_0, then f_n is convergent, up to subsequences.*

3. Setting

In this section we first explain our setting, namely what is an *embedded network* and what we mean by *Hamiltonian* on a network. Then we introduce the class of *Hamilton–Jacobi equations* on a network we are interested in, and specify the notions of solutions and subsolutions.

3A. *Embedded networks.* An *embedded network*, or *continuous graph*, is a subset $\Gamma \subset \mathbb{R}^N$ of the form

$$\Gamma = \bigcup_{\gamma \in \mathcal{E}} \gamma([0, 1]) \subset \mathbb{R}^N,$$

where $\mathcal{E}$ is a finite collection of regular (i.e., C^1 with nonvanishing derivative) simple oriented curves, called *arcs* of the network, that we assume, without any loss of generality, to be parametrized on [0, 1]. We denote by $\mathcal{E}^*$ the subset of arcs γ which are closed, namely with $\gamma(0) = \gamma(1)$.

Remark 3.1. Our setting can be easily extended to the case in which Γ is embedded in a Riemannian manifold (M, g), for example by means of Nash embedding theorem [1956]. Moreover, the results are independent of the chosen parametrizations of the arcs. In this regard, one could also choose a more intrinsic approach and consider arcs as 1-dimensional submanifolds, and the whole network as a stratified space. Hereafter we do not adopt this point of view.

Observe that on the support of any arc γ, we also consider the inverse parametrization defined as

$$\tilde{\gamma}(s) = \gamma(1 - s) \quad \text{for } s \in [0, 1].$$

We call $\tilde{\gamma}$ the *inverse arc* of γ. We assume

$$\gamma((0, 1)) \cap \gamma'([0, 1]) = \varnothing \quad \text{whenever } \gamma \ne \gamma' \text{ and } \gamma \ne \tilde{\gamma}'. \tag{2}$$

We call *vertices* the initial and terminal points of the arcs, and denote by $\boldsymbol{V}$ the sets of all such vertices. Note that (2) implies

$$\gamma((0, 1)) \cap \boldsymbol{V} = \varnothing \quad \text{for any } \gamma \in \mathcal{E}.$$

We assume that the network is connected; namely given two vertices there is a finite concatenation of arcs linking them.

The network Γ inherits a *geodesic distance*, denoted by d_Γ, from the Euclidean metric of $\mathbb{R}^N$. Hence, hereafter the notions of continuity and Lipschitz continuity, when referring to functions defined on Γ, must be understood with respect to such distance (which is indeed equivalent to the Euclidean one) and the induced topology.

We can also consider a *differential structure* on Γ by defining the tangent space at any $x \in \Gamma \setminus \boldsymbol{V}$ as

$$T_\Gamma(x) = \{\lambda\,\dot\gamma(t) \mid \lambda \in \mathbb{R},\ \gamma \in \mathcal{E},\ t \in (0,1) \text{ and } x = \gamma(t)\}$$

and the cotangent space $T^*_\Gamma(x)$ as the dual space $(T_\Gamma(x))^*$; namely, it is the set of linear functionals $p : T_\Gamma(x) \to \mathbb{R}$.

We will say that a function $f : \Gamma \to \mathbb{R}$ is of class $C^1(\Gamma \setminus V)$ if it is continuous in Γ and

$$t \mapsto f(\gamma(t)) \text{ is of class } C^1 \text{ in } (0,1) \text{ for any } \gamma \in \mathcal{E}.$$

For such a function we define $D_\Gamma f(x)$, where $x = \gamma(t_0)$ for some $\gamma \in \mathcal{E}$ and $t_0 \in (0,1)$, as the unique covector in $T^*_\Gamma(x)$ satisfying

$$(D_\Gamma f(x), \dot\gamma(t_0)) = \frac{d}{dt} f(\gamma(t))|_{t=t_0},$$

where $(\,\cdot\,,\cdot\,)$ denotes the pairing between covectors and vectors.

Notice that this definition is invariant for a change of parametrization from γ to $\tilde\gamma$.

We can associate to any continuous network Γ an abstract graph $\boldsymbol{X} = (\boldsymbol{V}, \boldsymbol{E})$ with the same vertices as the network and edges corresponding to the arcs. More precisely, we consider an abstract set $\boldsymbol{E}$ with a bijection

$$\Psi : \boldsymbol{E} \to \mathcal{E}. \tag{3}$$

This induces maps $o : \boldsymbol{E} \to \boldsymbol{V}$ and $^- : \boldsymbol{E} \to \boldsymbol{E}$ via

$$\mathrm{o}(e) = \Psi(e)(0) \quad \text{and} \quad \bar e = \Psi^{-1}(\widetilde{\Psi(e)})$$

satisfying the properties in the definition of graph. Intuitively, in the passage from the embedded network to the underlying abstract graph $\boldsymbol{X}$, the arcs become *immaterial* edges.

3B. ***Hamiltonians on networks.*** A Hamiltonian on a network Γ is a collection of Hamiltonians $\mathcal{H} = \{H_\gamma\}_{\gamma\in\mathcal{E}}$, where

$$H_\gamma : [0,1] \times \mathbb{R} \to \mathbb{R},$$
$$(s,p) \mapsto H_\gamma(s,p),$$

satisfies

$$H_{\tilde\gamma}(s,p) = H_\gamma(1-s,-p) \quad \text{for any } \gamma \in \mathcal{E}. \tag{4}$$

Notice that we are not assuming any periodicity on H_γ when γ is a closed curve.

We require any H_γ to be

(Hγ1) continuous in (s,p);

(Hγ2) coercive in p;

(Hγ3) quasiconvex in p; i.e., for every $a \in \mathbb{R}$ the set $\{p \in \mathbb{R} \mid H_\gamma(x, p) \le a\}$ is convex (provided it is nonempty). Moreover, we assume that

$$\mathrm{Int}\big(\{p \mid H_\gamma(x,p) \le a\}\big) = \{p \mid H_\gamma(x,p) < a\} \quad \text{for any } a \in \mathbb{R},$$

where $\mathrm{Int}(\,\cdot\,)$ denotes the interior of a set.

We point out that, throughout the paper, the term (sub-)solution to Hamilton–Jacobi equations involving the H_γ, must be understood in the viscosity sense; see for example [Bardi and Capuzzo-Dolcetta 1997; Barles 1994] for a comprehensive treatment of viscosity solutions theory.

We set for any $\gamma \in \mathcal{E}$

$$a_\gamma := \max_{s\in[0,1]} \min_{p\in\mathbb{R}} H_\gamma(s,p), \tag{5}$$

$$c_\gamma := \min\{a : H_\gamma = a \text{ admits periodic subsolutions}\}. \tag{6}$$

By periodic subsolution, we mean subsolution to the equation in (0, 1) taking the same value at the endpoints.

Remark 3.2. The definition of c_γ is indeed well-posed. In fact, given $\gamma \in \mathcal{E}$, because of the compactness of [0, 1], we can choose a large enough to have

$$H(s,0) \le a \quad \text{for any } s \in (0,1).$$

This shows that any constant function is a subsolution and, consequently, the set in the definition of c_γ is nonempty. It is also bounded from below since for $a < a_\gamma$ the corresponding equation does not admit subsolutions and, therefore, it does not admit periodic ones. Finally, by basic stability properties in viscosity solution theory, there exists a periodic subsolution at the level c_γ, which justifies the minimum appearing in the definition.

We will essentially use c_γ for $\gamma \in \mathcal{E}^*$, but in principle the definition and the above considerations hold for any γ.

We stress that

$$a_\gamma \le c_\gamma \quad \text{for any } \gamma \in \mathcal{E}.$$

We further define

$$a_0 := \max\Big\{\max_{\gamma\in\mathcal{E}\setminus\mathcal{E}^*} a_\gamma,\ \max_{\gamma\in\mathcal{E}^*} c_\gamma\Big\}. \tag{7}$$

We require a further condition:

(Hγ4) Given any $\gamma \in \mathcal{E}$ with $a_\gamma = a_0$, the map $s \mapsto \min_{p\in\mathbb{R}} H_\gamma(s,p)$ is constant in [0, 1].

Remark 3.3. The main role of (Hγ4) is to ensure uniqueness of solutions to the Dirichlet problem associated to the equation $H_\gamma = a_\gamma$, at least for the γ with $a_\gamma = a_0$. The uniqueness property for such kind of problems holds in general when the equation admits a strict subsolution, which is not the case at the level a_γ. The relevant consequence of condition (Hγ4) is that the family of subsolutions to $H_\gamma = a_\gamma$ reduces to a singleton, up to additive constants; see Proposition 5.3.

Finally, condition (Hγ4) is automatically satisfied if the H_γ are independent of the state variable.

3C. ***The eikonal Hamilton–Jacobi equation on networks.*** We define a notion of subsolution and solution to an equation of the form

$$\mathcal{H}(x, Du) = a \quad \text{on } \Gamma, \tag{$\mathcal{H}Ja$}$$

where $a \in \mathbb{R}$. This notation synthetically indicates the family (for γ varying in $\mathcal{E}$) of Hamilton–Jacobi equations

$$H_\gamma(s, (u \circ \gamma)') = a \quad \text{on } (0, 1). \tag{$HJ_\gamma a$}$$

We start by recalling some terminology of viscosity solutions theory.

Definition 3.4. Given a continuous function w in $[0, 1]$ and a function $\varphi \in C^1([0, 1]$, we say that:

- φ is *subtangent* to w at $s \in (0, 1)$ if

$$w = \varphi \quad \text{at } s \qquad \text{and} \qquad w \geq \varphi \quad \text{in } (s - \delta, s + \delta) \text{ for some } \delta > 0.$$

  The notion of *supertangent* is given by just replacing "$\geq$" by "$\leq$" in the above formula.

- φ is a *constrained subtangent* to w at 1 if

$$w = \varphi \quad \text{at } 1 \qquad \text{and} \qquad w \geq \varphi \quad \text{in } (1 - \delta, 1) \text{ for some } \delta > 0.$$

  A similar notion, with obvious adaptations, can be given at $t = 0$.

Definition 3.5. Given a continuous function w in $[0, 1]$ and a point $s_0 \in \{0, 1\}$, we say that w satisfies *the state-constraint boundary condition* for ($HJ_\gamma a$) at s_0 if

$$H_\gamma(s_0, \varphi'(s_0)) \geq a$$

for any φ that is a constrained C^1 subtangent to w at s_0.

Definition 3.6. We say that $u : \Gamma \to \mathbb{R}$ is a *subsolution* to ($\mathcal{H}Ja$) if

(i) it is continuous on Γ;

(ii) $s \mapsto u(\gamma(s))$ is a subsolution to ($HJ_\gamma a$) in $(0, 1)$ for any $\gamma \in \mathcal{E}$.

We say that u is *solution* to ($\mathcal{H}Ja$) if

(i) it is continuous;

(ii) $s \mapsto u(\gamma(s))$ is a solution of ($HJ_\gamma a$) in $(0, 1)$ for any $\gamma \in \mathcal{E}$;

(iii) for every vertex x there is at least one arc γ, having x as terminal point, such that $u(\gamma(s))$ satisfies the state-constraint boundary condition for ($HJ_\gamma a$) at $s = 1$.

Compare also this definition with the one in [Lions and Souganidis 2016]. As far as we know, the idea of imposing a supersolution condition on just one arc incident to a given vertex first appeared in [Schieborn and Camilli 2013].

We do not provide a notion of supersolution. This could be done straightforwardly but we will not need it in the remainder of the paper.

Definition 3.7. Given an open (in the relative topology) subset $\Gamma' \subset \Gamma$, we say that a continuous function $u : \Gamma \to \mathbb{R}$ is *solution to* $(\mathcal{H}Ja)$ *in* Γ' if for any $x \in \Gamma' \setminus \mathbf{V},\ x = \gamma(s_0)$ with $\gamma \in \mathcal{E},\ s_0 \in (0, 1)$, the usual viscosity solution condition holds true for $u \circ \gamma$ at s_0. If instead $x \in \Gamma' \cap \mathbf{V}$, we require condition (iii) in Definition 3.6 to hold.

Remark 3.8. The definition of (sub-)solutions on Γ requires $u \circ \gamma$ to be a (sub-)solution of the corresponding equation in (0, 1) on any arc γ. If, in particular γ is a closed curve, we must have in addition $u(\gamma(0)) = u(\gamma(1))$. This explains why on any arc $\gamma \in \mathcal{E}^*$ we are solely interested in periodic (sub-)solutions, namely (sub-)solutions in (0, 1) taking the same value at 0 and 1. This also explains the role of c_γ.

Remark 3.9. Let us point out that if the network is *augmented* by changing the status of a finite number of intermediate points of arcs in Γ, which become new vertices, then the notion of solution to $(\mathcal{H}Ja)$ is not affected. More specifically, if a function is a solution with respect to the original network, then it is still a solution for the augmented one; the converse property holds as well. This issue will be discussed more in detail in Remark 5.16.

Given a continuous function u defined in [0, 1], it is apparent that a C^1 function φ is supertangent (resp. subtangent) to u at $s_0 \in (0, 1)$ if and only if $\tilde{\varphi}(s) := \varphi(1 - s)$ is supertangent (resp. subtangent) to $s \mapsto u(1 - s)$ at $1 - s_0$. Taking into account (4), we derive the following result.

Proposition 3.10. *Given an arc γ, a function $u(s)$ is a subsolution (resp. solution) to $(HJ_\gamma a)$ if and only if $s \mapsto u(1 - s)$ is a subsolution (resp. solution) to the same equation with $H_{\tilde{\gamma}}$ in place of H_γ.*

It is not difficult to see that Lipschitz-continuity of subsolutions on any arc, coming from the coercivity condition in (Hγ2), implies Lipschitz-continuity in Γ with respect to the geodesic distance. We provide a proof in the Appendix for the reader's convenience.

Proposition 3.11. *The family of subsolutions to $(HJ_\gamma a)$, provided it is not empty, is equi-Lipschitz continuous on Γ with respect to the geodesic distance d_Γ.*

We derive from the previous result, plus basic properties of viscosity solutions, the existence of the maximal subsolution attaining a given value at a given point of the network.

Proposition 3.12. *Let a be such that the equation $(\mathcal{H}Ja)$ admits subsolution in Γ. Given $y \in \Gamma,\ \alpha \in \mathbb{R}$, the function*

$$w(x) = \max\{u(x) \mid \textit{subsolution to } (\mathcal{H}Ja) \textit{ with } u(y) = \alpha\}$$

is still a subsolution.

4. Main results and strategy of the proof

The remainder of the article consists of the proof of our results on existence, uniqueness and regularity of global (sub-)solutions to the eikonal Hamilton–Jacobi equation on Γ.

Main Theorem. *Let Γ be an embedded network (finite, connected, possibly including loops and more arcs connecting two vertices) and let $\boldsymbol{X} = (\boldsymbol{V}, \boldsymbol{E})$ be the underlying abstract graph. Let us consider a Hamiltonian $\mathcal{H} = \{H_\gamma\}_{\gamma \in \mathcal{E}}$ on the network, satisfying conditions* (Hγ1)–(Hγ4) *for any $\gamma \in \mathcal{E}$ and let a_0 denote the value defined in* (7).

I. *Global solutions*:

(i) (*existence*) *There exists a unique value $c = c(\mathcal{H}) \geq a_0$ — called the **Mañé critical value** — for which the equation $\mathcal{H}(x, Du) = c$ admits global solutions. In particular, these solutions are Lipschitz continuous on Γ.*

(ii) (*uniqueness*) *There exists a **uniqueness set** $\mathcal{A}_{\boldsymbol{X}} = \mathcal{A}_{\boldsymbol{X}}(\mathcal{H}) \subseteq \boldsymbol{V}$, called the (**projected**) **Aubry set** of $\mathcal{H}$, such that the following holds. Let $S_c : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}$ be the function defined in* (34); *then, given any **admissible trace** g on $\mathcal{A}_{\boldsymbol{X}}$, i.e., a function $g : \mathcal{A}_{\boldsymbol{X}} \to \mathbb{R}$ such that for every $x, y \in \mathcal{A}_{\boldsymbol{X}}$*

$$g(x) - g(y) \leq S_c(y, x),$$

there exists a unique global solution $u \in C(\Gamma, \mathbb{R})$ to $\mathcal{H}(x, Du) = c$ agreeing with g on $\mathcal{A}_{\boldsymbol{X}}$. Conversely, for any solution u to $\mathcal{H}(x, Du) = c$, the function $g = u|_{\mathcal{A}_X}$ gives rise to an admissible trace on $\mathcal{A}_{\boldsymbol{X}}$.

(iii) (*Hopf–Lax-type formula* 1) *Let $g : \mathcal{A}_{\boldsymbol{X}} \to \mathbb{R}$ be an admissible trace and $u \in C(\Gamma, \mathbb{R})$ the corresponding solution to $\mathcal{H}(x, Du) = c$. Then, on the support of any arc $\gamma \in \mathcal{E}$, u is given by*

$$u(\gamma(s)) = \min\{\boldsymbol{A}, \boldsymbol{B}\},$$

where

$$\boldsymbol{A} := \min\{g(y) + S_c(y, \gamma(0)) \mid y \in \mathcal{A}_{\boldsymbol{X}}\} + \int_0^s \sigma_c^+(t)\, dt,$$
$$\boldsymbol{B} := \min\{g(y) + S_c(y, \gamma(1)) \mid y \in \mathcal{A}_{\boldsymbol{X}}\} - \int_s^1 \sigma_c^-(t)\, dt,$$

with $s \in [0, 1]$ and σ_c^+, σ_c^- defined as in (8), (9) *with H_γ in place of H.*

(iv) (*Hopf–Lax-type formula* 2): *Let Γ' be a closed subset of Γ with*

$$\Gamma' \cap \gamma([0, 1]) \neq \varnothing \quad \text{for any } \gamma \text{ with } \Psi^{-1}(\gamma) \in \mathcal{A}_{\boldsymbol{X}}^*.$$

*For any **admissible trace** g on Γ', in the sense of* (65) *with c in place of a, there exists a unique solution $u \in C(\Gamma, \mathbb{R})$ to $\mathcal{H}(x, Du) = c$ agreeing with g on Γ', which is given by*

$$u(x) = \min\{g(y) + S_c^\Gamma(y, x) \mid y \in \Gamma'\},$$

where $S_c^\Gamma(\cdot, \cdot)$ denotes the intrinsic (semi-)distance defined in (63).

II. *Subsolutions*:

(i) (*maximal subsolutions*) *For $a \geq c$ and $y \in \Gamma$, the maximal subsolution to* ($\mathcal{H}Ja$) *taking an assigned value at y is a solution in $\Gamma \setminus \{y\}$.*

(ii) (*PDE characterization of the Aubry set*) *Let* $\mathcal{A}_\Gamma = \mathcal{A}_\Gamma(\mathcal{H}) \subset \Gamma$ *be the Aubry set on the network, as defined in* (49). *The maximal subsolution to* $(\mathcal{H}Jc)$ *taking a given value at a point* $y \in \Gamma$ *is a critical solution on the whole network if and only if* $y \in \mathcal{A}_\Gamma$.

(iii) (*regularity of critical subsolutions*) *Any subsolution* $v : \Gamma \to \mathbb{R}$ *to* $\mathcal{H}(x, Du) = c$ *is of class* $C^1(\Gamma \setminus \boldsymbol{V})$ *and they all possess the same differential on* $\mathcal{A}_\Gamma \setminus \boldsymbol{V}$. *More specifically, if* $x_0 \in \mathcal{A}_\Gamma$ *and* $x_0 = \gamma(s_0)$ *for some* $\gamma \in \mathcal{E}$ *and* $s_0 \in (0, 1)$, *then its differential at* x_0 *is uniquely determined by the relation*

$$(D_\Gamma v(x_0), \dot{\gamma}(s_0)) = \sigma_c^+(s_0),$$

where σ_c^+ *was defined in* (8), *and therefore*

$$v(\gamma(s)) = v(\gamma(0)) + \int_0^s \sigma_c^+(t)\, dt \quad \textit{for any } s \in [0, 1].$$

We infer from this that any pair of critical subsolutions differs by a constant on the support of γ.

(iv) (*existence of* C^1 *critical subsolutions*) *Given a function* $g : \boldsymbol{V} \to \mathbb{R}$ *such that*

$$g(x) - g(y) \le S_c(y, x) \quad \textit{for all } x, y \in \boldsymbol{V},$$

there exists a critical subsolution v *on* Γ, *with* $v = g$ *on* $\boldsymbol{V}$, *which is of class* C^1 *on* $\Gamma \setminus \boldsymbol{V}$. *In addition, there exists a critical subsolution* v *of class* $C^1(\Gamma \setminus \boldsymbol{V})$ *satisfying*

$$H_\gamma(s, Dv(\gamma(s))) < c$$

for all $s \in (0, 1)$ *and* $\gamma \in \mathcal{E}$ *with* $\gamma((0, 1)) \cap \mathcal{A}_\Gamma = \varnothing$.

(v) (*Hopf–Lax formula for maximal supercritical subsolutions* 1) *Let* $a > c$ *and* $\boldsymbol{V}' \subset \boldsymbol{V}$. *For any* $g : \boldsymbol{V}' \to \mathbb{R}$ *satisfying*

$$g(x) - g(y) \le S_a(y, x) \quad \textit{for all } x, y \in \boldsymbol{V}',$$

where $S_a(\cdot\,, \cdot)$ *was defined in* (34), *there exists a unique solution* u *to* $\mathcal{H}(x, Du) = a$ *in* $\Gamma \setminus \boldsymbol{V}'$ *agreeing with* g *on* $\boldsymbol{V}'$; *in addition,* u *is also a subsolution to* $\mathcal{H}(x, Du) = a$ *on the whole of* Γ. *In particular, on the support of any arc* $\gamma \in \mathcal{E}$, u *is given by*

$$u(\gamma(s)) = \min\{\boldsymbol{C},\ \boldsymbol{D}\},$$

where

$$\begin{aligned} \boldsymbol{C} &:= \tilde{g}(\gamma(0)) + \int_0^s \sigma_a^+(t)\, dt, \\ \boldsymbol{D} &:= \tilde{g}(\gamma(1)) - \int_s^1 \sigma_a^-(t)\, dt, \\ \tilde{g}(x) &:= \begin{cases} g(x) & \textit{if } x \in \boldsymbol{V}', \\ \min\{g(y) + S_a(y, x) \mid y \in \boldsymbol{V}'\} & \textit{if } x \notin \boldsymbol{V}', \end{cases} \end{aligned}$$

with $s \in [0, 1]$ *and* σ_a^+, σ_a^- *defined as in* (8), (9).

(vi) (*Hopf–Lax formula for maximal supercritical subsolutions* 2) *Let $a > c$ and Γ' be a closed subset of Γ. Let g be an admissible trace on Γ', in the sense of* (65), *then there exists a unique solution $u \in C(\Gamma, \mathbb{R})$ to* ($\mathcal{H}Ja$) *on $\Gamma \setminus \Gamma'$ agreeing with g on Γ', which is given by*

$$u(x) = \min\{g(y) + S_a^{\Gamma}(y, x) \mid y \in \Gamma'\},$$

where $S_a^{\Gamma}(\cdot, \cdot)$ denotes the intrinsic (*semi-*)*distance defined in* (63).

4A. ***Organization of the remaining sections and proof of the Main Theorem.*** For the sake of clarity, we provide here an outline of the forthcoming discussion and of the main steps involved in the proof.

In Section 5, we focus on the *local* problem on each arc of the network. Namely, for each $\gamma \in \mathcal{E}$ we study the existence of (sub-)solutions to the 1-dimensional eikonal Hamilton–Jacobi equation ($HJ_\gamma a$) with boundary conditions. In particular:

- We show that under suitable *admissibility conditions* on the boundary data, see (17), there exists a unique solution and we provide a representation formula (Proposition 5.5).
- We derive a characterization of condition (iii) in Definition 3.6 in terms of this representation formula (Proposition 5.6).

In Section 6 we concentrate on the global aspects of the problem:

- We introduce a *discrete functional equation* ($\mathcal{D}FEa$) on the abstract graph $\boldsymbol{X}$ and provide the corresponding notions of solutions and subsolutions. The crucial result linking solutions to this equation and solutions to ($\mathcal{D}FEa$) is proven in Proposition 6.2.
- In (30) we define the *Mañé critical value $c(\mathcal{H})$*. We first prove that this is the unique value for which solutions to the discrete functional equation may exist (Proposition 6.5), and then that the critical equation ($\mathcal{D}FEc$) indeed admits solutions (Theorem 6.16).
- In (39) and (40) we define the *Aubry set $\mathcal{A}^*_{\boldsymbol{X}}$* and the *projected Aubry set $\mathcal{A}_{\boldsymbol{X}}$*, which are nonempty (Lemma 6.20). We prove in Theorem 6.21 that $\mathcal{A}_{\boldsymbol{X}}$ is a uniqueness set and provide a Hopf–Lax-type representation formula for the solutions to ($\mathcal{D}FEc$) in terms of its values on $\mathcal{A}_{\boldsymbol{X}}$.

The supercritical case will be discussed in parallel to the critical one (see Propositions 6.3 and 6.6 and Theorem 6.23).

Finally, in Section 7 we switch our attention back to the immersed network:

- We prove in Theorem 7.1 that the notion of solution can be recovered in terms of maximal subsolution attaining a specific value at a given point.
- We introduce the analogue of the Aubry set on the network, we show in Theorem 7.5 that all critical subsolutions are of class C^1 on it and they all have the same differential on this set.
- We show the existence of C^1 critical subsolutions that are strict outside of the Aubry set (Theorem 7.6).
- We provide representation formulae and uniqueness results with traces that are not necessarily defined on vertices (Theorem 7.9).

For the reader's convenience, we provide here some references to the proof of each claim.

Proof of the Main Theorem.

(I) (i) Existence follows from Theorem 6.16 and Proposition 6.2; Lipschitz continuity follows from Proposition 3.11.
(ii) This part is obtained by combining Proposition 6.2 and Theorem 6.21.
(iii) This representation formula is proved in Proposition 5.5.
(iv) See Theorem 7.9(i).

(II) (i) See Theorem 7.1.
(ii) See Proposition 7.4.
(iii) See Theorem 7.5.
(iv) See Theorem 7.6.
(v) These results are obtained by combining Propositions 6.3 and 6.6 and Theorem 6.23 and using the representation formula in Proposition 5.5.
(vi) See Theorem 7.9(ii). □

5. Local part: the eikonal Hamilton–Jacobi equation with boundary conditions on arcs

In this section we focus on a single arc γ and study the family of equations $(HJ_\gamma a)$ in (0, 1), plus suitable boundary conditions. We assume

$$a \geq a_0 = \max\Big\{ \max_{\gamma\in\mathcal{E}\setminus\mathcal{E}^*} a_\gamma,\ \max_{\gamma\in\mathcal{E}^*} c_\gamma \Big\}.$$

Our aim is to find admissible conditions on boundary data at $s = 0$ and $s = 1$ to get solutions of the corresponding Dirichlet problem, to show uniqueness of such solutions and, finally, to provide a characterization of maximal subsolutions taking a given value at $s = 0$ via state-constraint boundary conditions.

We need specific results when γ is a closed curve because in this case we are solely interested in periodic (sub-)solutions, as explained in Remark 3.8. We address the issue in Section 5C. In Subsections 5A and 5B we will not distinguish between γ closed or not, and provide an unified presentation of the material.

The results are not new; we write down nevertheless the 1-dimensional representation formulae, which are easy to handle and allow a direct and simplified treatment of the matter. We recall that, due to coercivity and quasiconvexity assumptions, all subsolutions to $(HJ_\gamma a)$ are Lipschitz-continuous in [0, 1], and, in addition the notion of viscosity and a.e. subsolution are equivalent. Also notice that the subsolution property is not affected by addition of constants.

To ease notation, we write $H(s, p)$ instead of $H_\gamma(s, p)$, and accordingly we consider equation $(HJ_\gamma a)$ with H in place of H_γ. We recall that the assumptions (Hγ1)–(Hγ4) are in force.

5A. ***Setting of the local problem.*** We set, for $s \in [0, 1]$,

$$\sigma_a^+(s) = \max\{p \mid H(s, p) = a\}, \tag{8}$$

$$\sigma_a^-(s) = \min\{p \mid H(s, p) = a\}. \tag{9}$$

If $a > a_\gamma$, we have by (Hγ3)

$$(\sigma_a^-(s), \sigma_a^+(s)) = \{p \mid H(s, p) < a\} \quad \text{for } s \in [0, 1]. \tag{10}$$

We deduce from assumption (Hγ4) that if $a_\gamma = a_0$

$$\sigma_{a_\gamma}^+(s) = \sigma_{a_\gamma}^-(s) \quad \text{for any } s \in [0, 1]. \tag{11}$$

Proposition 5.1. *The functions $s \mapsto \sigma_a^+(s)$ and $s \mapsto \sigma_a^-(s)$ are continuous in* $[0, 1]$ *for any $a \geq a_\gamma$.*

Proof. It follows directly from the continuity and the coercivity of H that the function $s \mapsto \sigma_{a_\gamma}^+(s) = \sigma_{a_\gamma}^-(s)$ is continuous. If $a > a_\gamma$, the assertion follows from the fact that $\sigma_a^+(s)$ and $\sigma_a^-(s)$ are univocally determined for any s by the conditions $H(s, \sigma_a^+(s)) = H(s, \sigma_a^-(s)) = a$ and, respectively, $\sigma_a^+(s) > \sigma_{a_\gamma}^+(s)$ or $\sigma_a^-(s) < \sigma_{a_\gamma}^+(s)$. □

Notice that

$$u \text{ subsolution} \implies \sigma^-(s) \leq u'(s) \leq \sigma^+(s) \text{ for a.e. } s. \tag{12}$$

We introduce four relevant functions:

$$s \mapsto \int_0^s \sigma_a^+(t)\, dt, \tag{13}$$

$$s \mapsto \int_0^s \sigma_a^-(t)\, dt, \tag{14}$$

$$s \mapsto -\int_s^1 \sigma_a^-(t)\, dt, \tag{15}$$

$$s \mapsto -\int_s^1 \sigma_a^+(t)\, dt. \tag{16}$$

Remark 5.2. According to (12), the function in (13) is the maximal (sub-)solution to ($HJ_\gamma a$) vanishing at $s = 0$, and the one in (14) the minimal (sub-)solution vanishing at $s = 0$. Analogously, the function defined in (15) is the maximal (sub-)solution vanishing at $s = 1$, and the one in (16) the minimal (sub-)solution vanishing at $s = 1$. All of these functions are of class C^1 because of Proposition 5.1.

We remark that when we write *maximal (sub-)solution* and the like, we mean it is maximal in the class of subsolution to ($HJ_\gamma a$) with a given property and it is, in addition, a solution to the equation.

If $a = a_\gamma$, it follows from (11) that all of the above functions coincide up to an additive constant. We can state the following result.

Proposition 5.3. *The (sub-)solution to ($HJ_\gamma a$), with $a = a_\gamma$, is unique up to additive constants.*

From the properties of the solutions in (13) and (14), we directly derive a necessary condition (*admissibility condition*) that two boundary data at 0 and 1 must satisfy in order to correspond to the values at the endpoints of a subsolution to ($HJ_\gamma a$).

Lemma 5.4. *Assume that there is a subsolution to ($HJ_\gamma a$) taking the values α and β at* 0 *and* 1. *Then*

$$\int_0^1 \sigma_a^-(t)\, dt \leq \beta - \alpha \leq \int_0^1 \sigma_a^+(t)\, dt. \tag{17}$$

The above condition is actually also sufficient:

Proposition 5.5. *Given boundary data α, β satisfying* (17), *the function w,*

$$s \mapsto w(s) := \min\left\{\alpha + \int_0^s \sigma_a^+(t)\,dt,\ \beta - \int_s^1 \sigma_a^-(t)\,dt\right\}, \tag{18}$$

is the unique solution to $(HJ_\gamma a)$ taking the values α at $s = 0$ and β at $s = 1$.

The proof is in the Appendix.

5B. *Maximal subsolutions.* The main result of this section is:

Proposition 5.6. *Assume that w is a solution in* $(0, 1)$ *to $(HJ_\gamma a)$ for $a \geq a_\gamma$, continuously extended up to the boundary. If*

$$H(1, \varphi'(1)) \geq a \quad \textit{for any } C^1 \textit{ supertangent } \varphi \textit{ to } w \textit{ constrained to } [0, 1], \tag{19}$$

then w is the maximal (*sub-*)*solution taking the value $w(0)$ at* 0. *Namely*,

$$w(s) = w(0) + \int_0^s \sigma_a^+(t)\,dt \quad \textit{for } s \in [0, 1]. \tag{20}$$

Conversely, *if a solution w is of the form* (20), *then condition* (19) *holds true.*

The proof is in the Appendix.

We fix $s_0 \in (0, 1)$. By slightly generalizing the formulae provided in the previous result and arguing separately in the two subintervals $[0, s_0]$ and $[s_0, 1]$, we get:

Corollary 5.7. *Let $s_0 \in (0, 1)$. For any $\alpha \in \mathbb{R}$, the function*

$$s \mapsto \begin{cases} \alpha - \int_s^{s_0} \sigma_a^-(t)\,dt & \textit{for } s \leq s_0, \\ \alpha + \int_{s_0}^s \sigma_a^+(t)\,dt & \textit{for } s > s_0 \end{cases}$$

is the maximal subsolution to $(HJ_\gamma a)$ taking the value α at s_0. It is, *in addition*, *a solution in* $(0, 1) \setminus \{s_0\}$, *but the solution property fails at s_0, unless $a = a_\gamma$.*

Remark 5.8. In light of Proposition 3.10 and Remark 5.2, it is apparent that the maximal solution to $H(s, -u') = a$ vanishing at $s = 0$ is given by

$$s \mapsto -\int_{1-s}^1 \sigma_a^-(t)\,dt.$$

This function satisfies the state-constraint boundary condition at $s = 1$.

5C. *Closed arcs.* In this subsection we assume that γ is a closed curve. Keeping in mind Remark 3.8, we aim to show the existence of a periodic (sub-)solution for any a or, in other terms, that periodic boundary conditions at $s = 0$ and $s = 1$ are admissible in the sense of (17).

Recall that $a \geq a_0 \geq c_\gamma$. We derive further information in the case where $a = a_0 = c_\gamma$. We will exploit the existence of periodic subsolutions at the level c_γ in (0, 1), say, to fix ideas, vanishing at 0 and 1, as pointed out in Remark 3.2. These periodic subsolutions are *sandwiched* in between the function in (13) and the one in (14), according to Remark 5.2. We derive:

Lemma 5.9. *We have*

$$\int_0^1 \sigma_a^-(t)\,dt \le 0 \le \int_0^1 \sigma_a^+(t)\,dt, \tag{21}$$

and both the inequalities are strict if $a > c_\gamma$.

This, in view of (17), in turn implies:

Corollary 5.10. *There are periodic solutions to $(HJ_\gamma a)$ in $(0, 1)$.*

Moreover:

Proposition 5.11.
$$\min\left\{-\int_0^1 \sigma_{c_\gamma}^-(t)\,dt,\ \int_0^1 \sigma_{c_\gamma}^+(t)\,dt\right\} = 0.$$

The proof is in the Appendix.

From the previous result plus Proposition 5.6 and Remark 5.8, we derive the following.

Corollary 5.12. *Let $a = c_\gamma$ and $\alpha \in \mathbb{R}$; then, either the maximal solution to $H = a$ taking the value α at $s = 0$ or the maximal solution to $H(s, -u') = a$ taking the value α at $s = 0$ is periodic.*

In the final result of the section we provide a characterization for the maximal periodic subsolution taking a given value at $s_0 \in (0, 1)$. This corresponds, in the case of nonclosed arcs, to Corollary 5.7.

Corollary 5.13. *Let $s_0 \in (0, 1)$ and $\alpha \in \mathbb{R}$. We set*

$$\beta = \min\left\{-\int_0^{s_0} \sigma_a^-(t)\,dt,\ \int_{s_0}^1 \sigma_a^+(t)\,dt\right\}.$$

(i) *The maximal periodic subsolution to $(HJ_\gamma a)$ taking the value α at s_0, denoted by u, is uniquely determined by the condition of being solution of the equation in $(0, s_0)$ and $(s_0, 1)$ taking the values α at s_0 and $\alpha + \beta$ at 0 and 1.*

(ii) *If $\beta = -\int_0^{s_0} \sigma_a^-(t)\,dt$, then*

$$u(s) = \alpha - \int_s^{s_0} \sigma_a^-(t)\,dt \quad \textit{for } s \in [0, s_0]. \tag{22}$$

If instead $\beta = \int_{s_0}^1 \sigma^+ a(t)\,dt$, then

$$u(s) = \alpha + \int_{s_0}^{s} \sigma_a^+(t)\,dt \quad \textit{for } s \in [s_0, 1]. \tag{23}$$

The proof is in the Appendix.

5D. ***From local to global.*** The subsequent step in our analysis will be to transfer the Hamilton–Jacobi equation from Γ to the underlying graph $\boldsymbol{X}$, where it will take the form of a discrete functional equation. In doing this, the relevant information we derive from the above study is the value at $s = 1$ of the maximal solution to $H = a$ vanishing at $s = 0$. It is given, in accordance with Proposition 5.6, by

$$\int_0^1 \sigma_a^+(t)\,dt.$$

Therefore, if $\gamma = \Psi(e)$ and $a \geq a_\gamma$, we define

$$\sigma_a(e) := \int_0^1 \sigma_a^+(t)\,dt. \tag{24}$$

(recall that $a \geq a_0 \geq c_\gamma$).

Accordingly, we have

$$\sigma_a(-e) := -\int_0^1 \sigma_a^-(t)\,dt. \tag{25}$$

If e is a loop, or equivalently $\gamma = \Psi(e)$ a closed curve, we summarize the information gathered in Propositions 5.9 and 5.11 as follows:

Proposition 5.14. *If e is a loop then $\sigma_a(e) > 0$ for $a > c_\gamma$ and*

$$\min\{\sigma_{c_\gamma}(e),\ \sigma_{c_\gamma}(-e)\} = 0.$$

Moreover, we directly deduce from the definition of σ_a and (10) that:

Lemma 5.15. *The function*

$$a \mapsto \sigma_a(e)$$

is continuous and strictly increasing in $[a_\gamma, +\infty)$.

Remark 5.16. As already announced in Remark 3.9, we conclude this section with a remark on the invariance of the definition of solution to $(\mathcal{H}Ja)$ with respect to the addition of extra vertices to the network (*augmented network*). We discuss this issue in the case of a single extra vertex $x_0 = \gamma(s_0)$ for some $s_0 \in (0, 1)$ and γ a nonclosed arc.

We first prove that a solution u on Γ is also a solution for the augmented network. According to Proposition 5.5,

$$u(x_0) = \min\left\{u(\gamma(0)) + \int_0^{s_0} \sigma_a^+(t)\,dt,\ u(\gamma(1)) - \int_s^1 \sigma_a^-(t)\,dt\right\}.$$

If $u(x_0)$ is equal to the first term in the parentheses, then, by Proposition 5.6, u satisfies the state-constraint boundary condition with respect to the arc $\gamma|_{[0,s_0]}$, having the new vertex x_0 as terminal point. Whereas, if $u(x_0)$ equals the second term in the above formula, then the same property holds true for the arc $\tilde{\gamma}|_{[0,1-s_0]}$. This shows the claim.

To prove the converse, we start with a solution v to $(\mathcal{H}Ja)$ on the augmented network, with x_0 as the extra vertex, and consider the arcs $\gamma_1 = \gamma|_{[0,s_0]}$ and $\gamma_2 = \gamma|_{[s_0,1]}$, both parametrized on $[0, 1]$. The point is to show that the function $v \circ \gamma$ is a solution to $(HJ_\gamma a)$ in $(0, 1)$. It is apparently a subsolution in the whole interval and a solution in $(0, 1) \setminus \{s_0\}$. It also satisfies the state-constraint boundary condition at $s = 1$ either for the arc γ_1 or for $\tilde{\gamma}_2$. Since any subtangent to $v \circ \gamma$ at s_0 is a constrained subtangent at $s = 1$ for both γ_1 and $\tilde{\gamma}_2$, we deduce the supersolution property for $v \circ \gamma$ at s_0.

Arguing along the same lines, one can also check that the forthcoming notions of critical value and Aubry set are not affected by additions of new vertices.

6. Global part: the discrete functional equation on the abstract graph

In this section we push our analysis beyond the local existence of solutions to $(HJ_\gamma a)$ on each arc γ, and study the global existence of solutions to $(\mathcal{H}Ja)$ on the whole network Γ.

Let us start by noticing that if we consider $\boldsymbol{V}$, the set of vertices of Γ, it is easy to check that any solution w to $(\mathcal{H}Ja)$ has a well-defined *trace* $u = w|_V$ on $\boldsymbol{V}$, simply because of the continuity assumption. The following uniqueness result is straightforward. We provide a proof in the Appendix for reader's convenience.

Proposition 6.1. *Let u be a function defined on $\boldsymbol{V}$. Then there exists at most one solution to $(\mathcal{H}Ja)$ on Γ agreeing with u on $\boldsymbol{V}$.*

A converse property is by far more interesting, namely to find conditions on a function defined on $\boldsymbol{V}$ in order to (uniquely) extend it on the whole network as solution to $(\mathcal{H}Ja)$.

This issue — which is profoundly related to the global structure of the network — will be carefully addressed in this section.

More precisely, we study the problem of the admissibility, with respect to the fullref $(\mathcal{H}Ja)$, of a trace $g : \boldsymbol{V} \to \mathbb{R}$ defined on the global network and characterize all traces g that can be continuously extended to solutions to $(\mathcal{H}Ja)$ on the whole of Γ as solutions to an appropriate *discrete functional equation* on the underlying abstract graph $\boldsymbol{X} = (\boldsymbol{V}, \boldsymbol{E})$.

6A. ***The discrete functional equation.*** Given $a \geq a_0$, the cochain $\sigma_a \in \mathfrak{C}^1(\boldsymbol{X}, \mathbb{R})$ is defined as in (24), where $e = \Psi^{-1}(\gamma)$ and Ψ has been defined in (3).

If we recall the admissibility condition introduced in (17) plus (24) and (25), it is clear that the trace on $\boldsymbol{V}$ of a function $g : \Gamma \to \mathbb{R}$ admissible for the equations on any arc satisfies

$$-\sigma_a(-e) \leq \mathrm{d}g(e) = g(\mathrm{t}(e)) - g(\mathrm{o}(e)) \leq \sigma_a(e) \quad \text{for any } e \in \boldsymbol{E}, \tag{26}$$

which in particular implies

$$g(x) \leq \min_{e \in \boldsymbol{E}_x} \big(g(\mathrm{t}(e)) + \sigma_a(-e) \big) \quad \text{for } x \in \boldsymbol{V},$$

where $\boldsymbol{E}_x$ denotes the star centered at x, as defined in (1).

Inspired by this, we introduce the following *discrete functional equation*:

$$u(x) = \min_{e \in \boldsymbol{E}_x} \big(u(\mathrm{t}(e)) + \sigma_a(-e) \big) \quad \text{for } x \in \boldsymbol{V}. \tag{$\mathcal{D}FEa$}$$

Observe that the formulation of the discrete problem takes somehow into account the backward character of viscosity solutions.

A function v is a solution to $(\mathcal{D}FEa)$ in some subset $\boldsymbol{V}'$ of $\boldsymbol{V}$ if $(\mathcal{D}FEa)$ holds true with v in place of u and $x \in \boldsymbol{V}'$.

A function $u : \boldsymbol{V} \to \mathbb{R}$ is a *subsolution* to $(\mathcal{D}FEa)$ if

$$u(x) \leq \min_{e \in \boldsymbol{E}_x} \big(u(\mathrm{t}(e)) + \sigma_a(-e) \big) \quad \text{for } x \in \boldsymbol{V} \tag{27}$$

or, equivalently, if for each $e \in \boldsymbol{E}$ we have

$$du(e) \le \sigma_a(e), \tag{28}$$

which is equivalent to asking that $u(\mathrm{t}(e)) \le u(\mathrm{o}(e)) + \sigma_a(e)$ for each $e \in \boldsymbol{E}$.

A subsolution is qualified as *strict at* $x_0 \in \boldsymbol{V}$ if a strict inequality prevails in (27) when x is replaced by x_0. We say that u is *strict on a set* $A \subseteq \boldsymbol{V}$ if it is strict at every $x \in A$. We say that u is *strict* if it is strict everywhere on $\boldsymbol{V}$.

It is apparent that the property of being a solution or a subsolution is not affected by the addition of constants.

Our goal is to prove the existence of a solution to ($\mathcal{D}FEa$) (see Theorem 6.16). In fact, there is a crucial relation between the functional equation ($\mathcal{D}FEa$) and ($\mathcal{H}Ja$):

Proposition 6.2. *Given* $a \ge a_0$:

(i) *Any solution to* ($\mathcal{D}FEa$) *in* $\boldsymbol{V}$ *can be* (*uniquely*) *extended to a solution of* ($\mathcal{H}Ja$) *in* Γ; *conversely the trace on* $\boldsymbol{V}$ *of any solution of* ($\mathcal{H}Ja$) *in* Γ *is a solution to* ($\mathcal{D}FEa$).

(ii) *Any subsolution to* ($\mathcal{D}FEa$) *in* $\boldsymbol{V}$ *can be extended to a subsolution of* ($\mathcal{H}Ja$) *in* Γ; *conversely the trace on* $\boldsymbol{V}$ *of any subsolution of* ($\mathcal{H}Ja$) *in* Γ *is a subsolution to* ($\mathcal{D}FEa$).

Proof. Assume that u solves ($\mathcal{D}FEa$). Let x and y be two adjacent vertices, and e an edge with initial vertex x and final vertex y. We set $\gamma = \Psi(e)$ and consequently $\tilde{\gamma} = \Psi(-e)$; then $\gamma(0) = \tilde{\gamma}(1) = x$ and $\gamma(1) = \tilde{\gamma}(0) = y$. By the very definition of (sub-)solution to ($\mathcal{D}FEa$), we have

$$\begin{aligned} &u(\gamma(1)) - u(\gamma(0)) \le \sigma_a(e), \\ &u(\gamma(1)) - u(\gamma(0)) = u(\tilde{\gamma}(0)) - u(\tilde{\gamma}(1)) \ge -\sigma_a(-e). \end{aligned}$$

Taking into account (17), we derive that the values $u(\gamma(0))$ and $u(\gamma(1))$ are admissible for ($HJ_\gamma a$) in $(0, 1)$. We therefore deduce from Proposition 5.5 that there is a unique solution, say $w : [0, 1] \to \mathbb{R}$, to ($HJ_\gamma a$) taking precisely these values at the boundary. We define

$$v(z) = w(\gamma^{-1}(z)) \quad \text{for } z \in \gamma((0, 1)).$$

Since $\gamma((0, 1)) = \tilde{\gamma}((0, 1))$, one needs to check that this definition is well-posed, performing the same construction for $\tilde{\gamma}$, but this is a direct consequence of Proposition 3.10.

So far, we have successfully checked conditions (i) and (ii) in the definition of solution to ($\mathcal{H}Ja$) (see Definition 3.6). It is left to show (iii). Since u is a solution to ($\mathcal{D}FEa$), for any $x \in \boldsymbol{V}$ there is an edge e_0 with x as terminal vertex such that

$$u(x) - u(\mathrm{o}(e_0)) = \sigma_a(e_0).$$

Taking into account (24) and Proposition 5.6, for $\gamma = \Psi(e_0)$, we deduce that $v \circ \gamma$ actually satisfies the state-constraint boundary condition in (iii) with respect to ($HJ_\gamma a$).

Conversely, let u be a real function on $\boldsymbol{V}$ which is the trace on Γ of a solution to $(\mathcal{H}Ja)$. It follows from the compatibility condition (17), and the notations (24)–(25), that u is a subsolution to $(\mathcal{D}FEa)$; i.e.,

$$u(x) \le \min_{e \in \boldsymbol{E}_x} \big(u(\mathrm{t}(e)) + \sigma_a(-e) \big) \quad \text{for } x \in \boldsymbol{V}. \tag{29}$$

In order to show that it is a solution to $(\mathcal{D}FEa)$, we need to prove that equality holds in (29) for every $x \in \boldsymbol{V}$. In fact, since u is the trace of a solution to $(\mathcal{H}Ja)$, it follows from condition (iii) in Definition 3.6, that for every vertex x there is at least one arc γ having x as terminal point such that $u(\gamma(s))$ satisfies the state-constraint boundary condition for $(HJ_\gamma a)$ at $s = 1$. In particular, in light of Proposition 5.6, see (24), this implies that there exists e with $\mathrm{t}(e) = x$, or in other terms $-e \in \boldsymbol{E}_x$, such that

$$u(x) - u(\mathrm{o}(e)) = \sigma_a(e)$$

or equivalently

$$u(x) = u(\mathrm{t}(-e)) + \sigma_a(e).$$

Hence, equality holds in (29), and this completes the proof of item (i). Item (ii) can be proven arguing along the same lines. □

The same argument as in the above proof allows also showing the following:

Proposition 6.3. *Given $a \ge a_0$ and $\boldsymbol{V}' \subset \boldsymbol{V}$, a function $u : \boldsymbol{V} \to \mathbb{R}$ which is a subsolution to $(\mathcal{D}FEa)$ in $\boldsymbol{V}$ and solution in $\boldsymbol{V} \setminus \boldsymbol{V}'$ can be (uniquely) extended to a function $v : \Gamma \to \mathbb{R}$ which is a subsolution of $(\mathcal{H}Ja)$ in Γ and a solution in $\Gamma \setminus \boldsymbol{V}'$. Conversely, the trace on $\boldsymbol{V}$ of a function $v : \Gamma \to \mathbb{R}$ which is a subsolution to $(\mathcal{H}Ja)$ in Γ and a solution in $\Gamma \setminus \boldsymbol{V}'$ is a subsolution to $(\mathcal{D}FEa)$ in $\boldsymbol{V}$ and a solution in $\boldsymbol{V} \setminus \boldsymbol{V}'$.*

6B. *Existence of solutions to (𝒟FEa) and critical value.* We want to introduce a notion of *critical value* for $(\mathcal{D}FEa)$ and prove the existence of solutions.

Let us start by proving the following stability properties of solutions and subsolutions.

Proposition 6.4. (i) *Let a_n be a sequence in $\mathbb{R}$ converging to some a. Let u_n be subsolution to $(\mathcal{D}FEa_n)$ for every n, with $u_n(x_0)$ bounded for some $x_0 \in \boldsymbol{V}$; then u_n converges, up to subsequences, to a subsolution to $(\mathcal{D}FEa)$.*

(ii) *Let v_n be a sequence of solutions to $(\mathcal{D}FEa)$ for some $a \in \mathbb{R}$, with $v_n(x_0)$ bounded for some $x_0 \in \boldsymbol{V}$; then v_n converges, up to a subsequence, to a solution to $(\mathcal{D}FEa)$.*

Proof. Owing to the definition of subsolution and Lemma 5.15, we see that

$$\langle \mathrm{d}u_n, e \rangle \le \sigma_b(e) \quad \text{for every } e \in \boldsymbol{E},$$

where $b = \sup a_n$. This implies that the $\mathrm{d}u_n$ are equibounded. We therefore get, exploiting the boundedness assumption on x_0 and Proposition 2.1(ii), that u_n is convergent, up to subsequences, to some u. By Lemma 5.15 we have

$$u(\mathrm{t}(e)) - u(\mathrm{o}(e)) - \sigma_a(e) = \lim_n \big(u_n(\mathrm{t}(e)) - u_n(\mathrm{o}(e)) - \sigma_{a_n}(e) \big) \le 0$$

for any e, showing that u is a subsolution to $(\mathcal{D}FEa)$.

Let now v_n be a sequence of solutions to ($\mathcal{D}FEa$); because of the previous point, v_n converges, up to subsequences, to a subsolution v of the same equation. It is left to show that v is indeed a solution. Given $x \in \boldsymbol{V}$, we find $e_n \in \boldsymbol{E}_x$ with

$$v_n(\mathrm{t}(e_n)) - v_n(x) - \sigma_a(-e_n) = 0.$$

Since the edges are finite, we deduce that there exists $e_0 \in \boldsymbol{E}_x$ such that

$$e_n = e_0 \quad \text{for infinitely many } n.$$

Up to extracting a subsequence, passing to the limit as n goes to infinity, we obtain

$$v(\mathrm{t}(e_0)) - v(x) - \sigma_a(-e_0) = 0. \qquad \square$$

We define the *critical value* for ($\mathcal{D}FEa$) (also called the *Mañé critical value*) as

$$c = c(\mathcal{H}) := \min\{a \geq a_0 \mid (\mathcal{D}FEa) \text{ admits subsolutions}\}. \tag{30}$$

First of all, notice that it is well-defined. In fact, because of the coercivity of the H_γ, we know σ_a is strictly positive for every e, when a is large enough, so that any constant function is a subsolution to ($\mathcal{D}FEa$). This shows that c is finite. Note the minimum in the definition of c is justified by Proposition 6.4, showing the existence of critical subsolutions (namely, subsolutions to ($\mathcal{D}FEa$) with $a = c$).

The relevance of the critical value is apparent from the following result.

Proposition 6.5. *If there exists a solution to* ($\mathcal{D}FEa$), *then* $a = c$.

Proof. Clearly $a \geq c$, since every solution is also a subsolution. If $a > c$, then there exists a strict subsolution u to ($\mathcal{D}FEa$). Let us assume, by contradiction, that there exists also a solution v. Let x_0 be point at which $u - v$ achieves its maximum; then

$$v(x_0) - v(\mathrm{t}(e)) \leq u(x_0) - u(\mathrm{t}(e)) \quad \text{for any } e \in \boldsymbol{E}_{x_0}. \tag{31}$$

By the very definition of solution applied to v, there is $e_0 \in \boldsymbol{E}_{x_0}$ such that

$$v(x_0) = v(\mathrm{t}(e_0)) + \sigma_a(-e_0).$$

We derive, taking into account (31),

$$u(x_0) \geq u(\mathrm{t}(e_0)) + \sigma_a(-e_0),$$

which is in contrast with the very definition of strict subsolution. $\square$

We further deduce a uniqueness result in the supercritical case.

Proposition 6.6. *Let* $a > c$ *and* $\boldsymbol{V}' \subset \boldsymbol{V}$. *For any given function* u *defined on* $\boldsymbol{V}'$ *there is at most one solution* v *of* ($\mathcal{D}FEa$) *in* $\boldsymbol{V} \setminus \boldsymbol{V}'$ *agreeing with* u *on* $\boldsymbol{V}'$.

Proof. Assume by contradiction that there are two distinct solutions u_1 and u_2 both satisfying the statement. Since $a > c$, we know that there is a strict subsolution w to ($\mathcal{D}FEa$). Therefore, given $\lambda \in (0, 1)$ we have

$$\lambda\, w(x) + (1-\lambda)\, u_1(x) < \min_{e \in \boldsymbol{E}_x} \big(\lambda\, w(\mathrm{t}(e)) + (1-\lambda)\, u_1(\mathrm{t}(e)) \, + \, \sigma_a(-e)\big) \tag{32}$$

for any $x \in \boldsymbol{V} \backslash \boldsymbol{V}'$. Up to interchanging the roles of u_1 and u_2, we can assume that $\max_{\boldsymbol{V}}(u_1 - u_2) > 0$, so that any maximizer is outside $\boldsymbol{V}'$. For λ sufficiently close to 0, we still have that $[\lambda\, w + (1-\lambda)\, u_1] - u_2$ achieves its maximum in $\boldsymbol{V} \backslash \boldsymbol{V}'$. Let x_0 be one of these points of maximum; then, for every $e \in \boldsymbol{E}_{x_0}$ we have

$$[\lambda\, w(x_0) + (1-\lambda)\, u_1(x_0)] - u_2(x_0) \geq [\lambda\, w(\mathrm{t}(e)) + (1-\lambda)\, u_1(\mathrm{t}(e))] - u_2(\mathrm{t}(e))$$

or

$$u_2(x_0) \leq u_2(\mathrm{t}(e)) + \lambda\, w(x_0) + (1-\lambda)\, u_1(x_0) - \lambda\, w(\mathrm{t}(e)) - (1-\lambda)\, u_1(\mathrm{t}(e)).$$

Using (32) we can deduce

$$u_2(x_0) < \min_{e \in \boldsymbol{E}_{x_0}} \big(u_2(\mathrm{t}(e)) - \sigma_a(-e)\big)$$

in contrast with $x_0 \notin \boldsymbol{V}'$ and u_2 being solution to ($\mathcal{D}FEa$) in $\boldsymbol{V} \backslash \boldsymbol{V}'$. □

Given $a \geq a_0$, we define for any path $\xi = (e_1, \ldots, e_M) = (e_i)_{i=1}^M$,

$$\sigma_a(\xi) = \sum_{i=1}^{M} \sigma_a(e_i), \tag{33}$$

and

$$S_a(x, y) := \inf\{\sigma_a(\xi) \mid \xi \text{ is a path linking } x \text{ to } y\}. \tag{34}$$

The following triangle inequality is a direct consequence of the definition:

$$S_a(x, y) \leq S_a(x, z) + S_a(z, y) \quad \text{for any } x, y, z \text{ in } \boldsymbol{V}. \tag{35}$$

The next result starts unveiling the major role of cycles in the forthcoming analysis.

Lemma 6.7. *$S_a \not\equiv -\infty$ if and only if*

$$\sigma_a(\xi) \geq 0 \quad \textit{for any cycle } \xi,$$

which is equivalent to saying that $S_a(x, x) \geq 0$ for any $x \in \boldsymbol{V}$.

Proof. If $\sigma_a(\xi) < 0$ for some cycle ξ, then going through it several times, we deduce that $S_a \equiv -\infty$. Conversely, if $\sigma_a(\xi) \geq 0$ for any cycle ξ, then

$$S_a(x, x) \geq 0 \quad \text{for any } x \in \boldsymbol{V}$$

and therefore $S_a \not\equiv -\infty$. □

From the very definition of subsolution we derive the following result.

Proposition 6.8. *A function u is a subsolution to ($\mathcal{D}FEa$) if and only if*

$$u(x) - u(y) \leq S_a(y, x) \quad \textit{for any } x, y \in \boldsymbol{V}.$$

Proof. It follows easily from the definitions of subsolution in (28) and σ_a in (33) that

$$u(x) - u(y) \leq \sigma_a(\xi) \quad \text{for any path } \xi \text{ linking } y \text{ to } x.$$

Taking the minimum over all such paths, we get the inequality in the statement. The converse is trivial, observing that

$$S_a(\mathrm{o}(e), \mathrm{t}(e)) \le \sigma_a(e) \quad \text{for every } e \in \boldsymbol{E}. \qquad \square$$

The previous result implies:

Corollary 6.9. *If $a \ge c$ then $S_a \not\equiv -\infty$.*

Moreover:

Corollary 6.10. *Given $a \ge c$ and x, y in $\boldsymbol{V}$, there exists a simple path η with $\mathrm{o}(\eta) = x$ and $\mathrm{t}(\eta) = y$ such that $\sigma_a(\eta) = S_a(x, y)$.*

Proof. Let $\xi = (e_i)_{i=1}^M$ be any path linking x to y. If ξ is not simple there are indices $k > j$ such that $\mathrm{t}(e_i) = \mathrm{t}(e_j)$. We assume, to ease notation, that $k < M$; the case $k = M$ can be treated with straightforward modifications.

We have that $(e_i)_{i=j+1}^k$ is a cycle and the paths $(e_i)_{i=1}^j$ and $(e_i)_{i=k+1}^M$ are concatenated. We get, according to Lemma 6.7, that

$$\sigma_a(\xi) = \sigma_a((e_i)_{i=1}^j) + \sigma_a((e_i)_{i=j+1}^k) + \sigma_a((e_i)_{i=k+1}^M) \ge \sigma_a((e_i)_{i=1}^j) + \sigma_a((e_i)_{i=k+1}^M)$$

and $(e_i)_{i=1}^j \cup (e_i)_{i=k+1}^M$ is still a path linking x to y. By iterating the above procedure, we remove all cycles properly contained in ξ and end up with a simple curve ξ_0 with $\mathrm{o}(\xi_0) = x$, $\mathrm{t}(\xi_0) = y$ and $\sigma_a(\xi_0) \le \sigma_a(\xi)$. This shows that $S_a(x, y)$ can be realized as the infimum of simple paths from x to y. Since there are finitely many such paths, we get the assertion. $\square$

The condition in Corollary 6.9 is actually necessary and sufficient, as shown by the next result. In the proof we will use a form of the basic *Bellman optimality principle* adapted to our frame. It can be stated as follows: if $\xi = (e_i)_{i=1}^M$ is a path with

$$\sigma_a(\xi) = S_a(\mathrm{o}(e), \mathrm{t}(e))$$

and $1 \le j < k \le M$, then $\eta := (e_i)_{i=j}^k$ satisfies $\sigma_a(\eta) = S_a(\mathrm{o}(e_j), \mathrm{t}(e_k))$.

Proposition 6.11. *Assume $S_a \not\equiv -\infty$. Given $y \in \boldsymbol{V}$, the function $u = S_a(y, \cdot)$ is a solution to ($\mathcal{D}$FEa) in $\boldsymbol{V} \setminus \{y\}$ and a subsolution to ($\mathcal{D}$FEa) in $\boldsymbol{V}$.*

Proof. The subsolution property comes from Proposition 6.8 and the triangle inequality (35). We proceed by showing that u is a solution in $\boldsymbol{V} \setminus \{y\}$. Let $x \ne y$; then, by Corollary 6.10, there is a path $\xi = (e_i)_{i=1}^M$ linking y to x with

$$\sigma_a(\xi) = S_a(y, x).$$

By the Bellman optimality principle, the path $\eta := (e_i)_{i=1}^{M-1}$ satisfies

$$\sigma_a(\eta) = S_a(y, \mathrm{t}(\eta)) = u(\mathrm{t}(\eta)).$$

Consequently

$$u(x) = \sigma_a(\eta) + \sigma_a(e_M) = u(\mathrm{t}(\eta)) + \sigma_a(e_M)$$

with $-e_M \in \boldsymbol{E}_x$. Hence

$$u(x) - u(\mathrm{t}(-e_M)) = u(x) - u(\mathrm{t}(\eta)) = \sigma_a(e_M). \qquad \square$$

Using Proposition 6.8 and the triangle inequality (35), we also obtain

Corollary 6.12. *The function*

$$x \mapsto -S_c(x, y)$$

is a critical subsolution for any fixed $y \in \boldsymbol{V}$.

Combining Corollary 6.9 and Proposition 6.11 we get:

Corollary 6.13. $S_a \not\equiv -\infty$ *if and only if* $a \geq c$.

We further have:

Proposition 6.14. *Given* $y \in \boldsymbol{V}$, *the function* $x \mapsto S_a(y, x)$ *is a solution to* ($\mathcal{D}FEa$) *if and only if there exists a cycle* ξ *incident on* y *with* $\sigma_a(\xi) = 0$.

Proof. ($\Longrightarrow$) We will prove in Proposition 6.15 a more general property, namely that if the equation ($\mathcal{D}FEa$) admits a solution, then there is a cycle ξ with $\sigma_a(\xi) = 0$.

($\Longleftarrow$) Assume the existence of a cycle, say $\xi = (e_i)_{i=1}^M$, with $\sigma_a(\xi) = 0$ incident on y. Up to relabelling the e_i, we can set $y = \mathrm{o}(\xi) = \mathrm{t}(\xi)$. We claim that $u := S_a(y, \cdot)$ is a solution on the whole of $\boldsymbol{V}$. By Proposition 6.11, it is enough to prove the assertion at y. We have

$$0 \leq S_a(y, y) = u(y) \leq \sigma_a(e_M) + S_a(y, \mathrm{o}(e_M)) \leq \sigma_a(\xi),$$

and since $\sigma_a(\xi) = 0$, all the inequalities in the above formula must indeed be equalities; in particular

$$u(y) - u(\mathrm{t}(-e_M)) - \sigma_a(e_M) = u(y) - S_a(y, \mathrm{o}(e_M)) - \sigma_a(e_M) = 0$$

with $-e_M \in \boldsymbol{E}_y$. This proves the claim. □

As announced, we complete the above proof by showing:

Proposition 6.15. *If the equation* ($\mathcal{D}FEa$) *admits a solution, then there is a cycle* ξ *with* $\sigma_a(\xi) = 0$.

Proof. Let us assume that v is a solution to ($\mathcal{D}FEa$). Take any $x \in \boldsymbol{V}$; by the definition of solution, we can find an edge e with terminal vertex x such that

$$v(x) - v(\mathrm{o}(e)) = \sigma_a(e).$$

By iterating backward the procedure, we can construct for any M a path $\xi = (e_i)_{i=1}^M$ such that

$$v(\mathrm{t}(e_j)) - v(\mathrm{o}(e_k)) = \sigma_a((e_i)_{i=k}^j) \quad \text{for any } j \geq k. \tag{36}$$

Since the graph is finite, taking M large enough, we have that for suitable indices $j > k$, the path $(e_i)_{i=k}^j$ is a cycle, so that $v(\mathrm{t}(e_j)) - v(\mathrm{o}(e_k)) = 0$, and the relation (36) provides the assertion. □

The argument of the next proof is reminiscent of the one used for the existence of critical solutions of Hamilton–Jacobi equations in compact manifolds; see [Fathi and Siconolfi 2005].

Theorem 6.16. *The critical equation* ($\mathcal{D}FEc$) *admits solutions.*

Proof. We break up the argument according to whether $c = a_0$ or $c > a_0$. Let us first discuss the first instance. If in addition $c = a_\gamma$ for some arc γ, and we set $e = \Psi^{-1}(\gamma)$, then we get from (11), (24), (25) that

$$\sigma_c(e \cup (-e)) = 0.$$

If instead $a_0 = c_\gamma$ for some closed arc γ of the network, then $e = \Psi^{-1}(\gamma)$ is a loop and we obtain, by Proposition 5.14,

$$\sigma_c(e) = 0 \quad \text{or} \quad \sigma_c(-e) = 0.$$

In both cases, we infer the existence of a critical solution in light of Proposition 6.14.

We proceed considering the case $c > a_0$. Let us assume by contradiction that there are no critical solutions. For any $y \in \boldsymbol{V}$, setting $u_y = S_c(y, \cdot)$, we can therefore find by Proposition 6.11 a positive constant δ_y with

$$\max_{e \in \boldsymbol{E}_y} \big(u_y(y) - u_y(\mathrm{t}(e)) - \sigma_c(-e)\big) = -\delta_y. \tag{37}$$

We define $u = \sum_y \lambda_y\, u_y$, where the λ_y are positive coefficients summing to 1, and set

$$\delta = \min_y \lambda_y\, \delta_y.$$

Exploiting that all the u_y are subsolutions on the whole of $\boldsymbol{V}$ and using (37), we conclude that for any $e \in \boldsymbol{E}$

$$\begin{aligned} u(\mathrm{t}(e)) - u(\mathrm{o}(e)) - \sigma_c(e) &= \sum_{y \neq \mathrm{t}(e)} \lambda_y \big(u_y(\mathrm{t}(e)) - u_y(\mathrm{o}(e)) - \sigma_c(e)\big) + \lambda_{\mathrm{t}(e)} \big(u_{\mathrm{t}(e)}(\mathrm{t}(e)) - u_{\mathrm{t}(e)}(\mathrm{o}(e)) - \sigma_c(e)\big) \\ &\leq -\lambda_{\mathrm{t}(e)}\, \delta_{\mathrm{t}(e)} \leq -\delta. \end{aligned} \tag{38}$$

Owing to Lemma 5.15 and the fact that $c > a_0$, there is $a_0 < b < c$ with

$$\sigma_b(e) > \sigma_c(e) - \delta \quad \text{for every } e \in \boldsymbol{E};$$

then we deduce from (38) that

$$u(\mathrm{t}(e)) - u(\mathrm{o}(e)) - \sigma_b(e) \leq 0 \quad \text{for every } e.$$

This proves that u is a subsolution to ($\mathcal{D}FEa$) with $a = b$, which is impossible because $b < c$. Therefore the maximum in (37) must be 0 for some y_0, which in turn implies that $S_c(y_0, \cdot)$ is a critical solution, as was claimed. □

Remark 6.17. Let u be a solution to ($\mathcal{D}FEc$). Let e be a loop with $\mathrm{o}(e) = \mathrm{t}(e) = x$, and $\gamma = \Psi(e)$ is hence a closed curve. If $c < c_\gamma$, then, according to Proposition 5.14

$$0 = u(\mathrm{o}(e)) - u(\mathrm{t}(e)) < \sigma_c(e), \quad 0 = u(\mathrm{o}(-e)) - u(\mathrm{t}(-e)) < \sigma_c(-e),$$

which shows that neither e nor $-e$ realizes

$$\min_{e \in \boldsymbol{E}_x} \big(u(\mathrm{t}(e)) + \sigma_a(-e)\big).$$

This in turn implies that the edge e, and consequently $-e$, can be removed from the edges of $\boldsymbol{X}$ without affecting the status of solution for u or any other critical solution.

Things are different if $c = c_\gamma$ because in this case, see Proposition 5.14,

$$0 = \min\{\sigma_c(e), \sigma_c(-e)\} = u(\mathrm{o}(e)) - u(\mathrm{t}(e)) = u(\mathrm{o}(-e)) - u(\mathrm{t}(-e)).$$

6C. ***The Aubry set and some structural properties of solutions.*** Inspired by what was discussed in the previous subsection, we introduce the following definition.

Definition 6.18. The *Aubry set* is defined as

$$\mathcal{A}^*_{\boldsymbol{X}} = \mathcal{A}^*_{\boldsymbol{X}}(\mathcal{H}) = \{e \in \boldsymbol{E} \mid \text{belonging to some cycle with } \sigma_c(\xi) = 0\}. \tag{39}$$

The *projected Aubry set* is given by

$$\mathcal{A}_{\boldsymbol{X}} = \mathcal{A}_{\boldsymbol{X}}(\mathcal{H}) = \{y \in \boldsymbol{V} \mid \exists\, \xi \text{ cycle incident on } y \text{ with } \sigma_c(\xi) = 0\}. \tag{40}$$

The projected Aubry set is partitioned into *static classes*, defined as the equivalence classes with respect to the relation

$$S_c(x, y) + S_c(y, x) = 0.$$

Equivalently x and y belong to the same static class if there is a cycle ξ with $\sigma_c(\xi) = 0$ incident on both of them; in particular, the whole cycle ξ is then contained in this static class.

Remark 6.19. Clearly, $x \in \mathcal{A}_{\boldsymbol{X}}$ if and only if $x = \mathrm{o}(e) = \mathrm{t}(e')$ for some e, e' in $\mathcal{A}^*_{\boldsymbol{X}}$; moreover, if $e \in \mathcal{A}^*_{\boldsymbol{X}}$, then $\mathrm{o}(e)$ and $\mathrm{t}(e)$ belong to $\mathcal{A}_{\boldsymbol{X}}$. The converse of this last property is not true because, for instance, if $e \in \mathcal{A}^*_{\boldsymbol{X}}$ then $-e$ might not belong to $\mathcal{A}^*_{\boldsymbol{X}}$. It is also possible to have a pair of adjacent vertices belonging to different static classes of $\mathcal{A}_{\boldsymbol{X}}$ linked by an edge not in $\mathcal{A}^*_{\boldsymbol{X}}$, or even vertices of the same static classes linked by multiple edges not all belonging to $\mathcal{A}^*_{\boldsymbol{X}}$.

We immediately derive from Proposition 6.15 and Theorem 6.16 the following result.

Lemma 6.20. *The Aubry sets are nonempty. Moreover*,

$$\mathcal{A}_{\boldsymbol{X}} = \{y \in \boldsymbol{V} \mid S_c(y, y) = 0\} = \{y \in \boldsymbol{V} \mid S_c(y, \cdot\,) \textit{ is a solution to } (\mathcal{D}FEc)\}.$$

We have a structural result on critical solutions. By admissible trace g on $\boldsymbol{V}' \subset \boldsymbol{V}$ (for the critical equation), we mean a function satisfying

$$g(x) - g(y) \le S_c(y, x) \quad \text{for any } x,\ y \text{ in } \boldsymbol{V}'. \tag{41}$$

Theorem 6.21. *Given an admissible trace g on $\mathcal{A}_{\boldsymbol{X}}$, the unique solution to* ($\mathcal{D}FEc$) *taking the value g on $\mathcal{A}_{\boldsymbol{X}}$ is*

$$v(x) := \min\{g(y) + S_c(y, x) \mid y \in \mathcal{A}_{\boldsymbol{X}}\}. \tag{42}$$

In particular, $\mathcal{A}_{\boldsymbol{X}}$ *represents a uniqueness set for the equation.*

Proof. Taking into account (41) and the fact that $S_c(y, y) = 0$ for any $y \in \mathcal{A}_{\boldsymbol{X}}$, we deduce that g and v coincide on $\mathcal{A}_{\boldsymbol{X}}$. The function v is a critical solution, since it is the pointwise minimum of a finite family of solutions. This property can be easily derived from the definition of solution.

Assume now that w is another solution agreeing with g on $\mathcal{A}_X$. Given any $x \in \boldsymbol{V}$, we construct, arguing as in Proposition 6.15, a path $\xi = (e_i)_{i=1}^M$ with $\mathrm{t}(\xi) = x$ and such that

$$w(\mathrm{t}(e_j)) - w(\mathrm{o}(e_k)) = \sigma_c((e_i)_{i=k}^j) \quad \text{for any } j \geq k.$$

If M is sufficiently large, there must exist $j_0 \geq k_0$ such that $(e_i)_{i=k_0}^{j_0}$ is a cycle. We deduce that there are $y \in \mathcal{A}_X$ and a path η linking y to x with

$$w(x) = w(y) + \sigma_c(\eta) \geq g(y) + S_c(y, x) \geq v(x).$$

Since the converse inequality holds true by Proposition 6.8, we get $w(x) = v(x)$. ☐

We record for later use an immediate consequence of the above result:

Corollary 6.22. *Given $\boldsymbol{V}' \subset \mathcal{A}_X$, and an admissible trace g on it, the function*

$$v(x) := \min\{g(y) + S_c(y, x) \mid y \in \boldsymbol{V}'\} \tag{43}$$

is the maximal solution to ($\mathcal{D}FEc$) *taking the value g on $\boldsymbol{V}'$.*

We can also derive a representation formula for solutions at $a > c$ in some subset of $\boldsymbol{V}$. To help in understanding the next statement, we recall that $S_a(x, x) > 0$ for any $x \in \boldsymbol{V}$ whenever $a > c$.

Theorem 6.23. *Let $a > c$ and $\boldsymbol{V}' \subset \boldsymbol{V}$. Let g be a function defined on $\boldsymbol{V}'$ satisfying* (41) *with S_a in place of S_c. Then the function*

$$v(x) = \begin{cases} g(x) & \text{if } x \in \boldsymbol{V}', \\ \min\{g(y) + S_a(y, x) \mid y \in \boldsymbol{V}'\} & \text{if } x \notin \boldsymbol{V}' \end{cases}$$

is the unique solution to ($\mathcal{D}FEa$) *in $\boldsymbol{V} \setminus \boldsymbol{V}'$ agreeing with g on $\boldsymbol{V}'$. It is in addition a subsolution on the whole of $\boldsymbol{V}$.*

Proof. We claim that

$$v(z) - v(x) \leq S_a(x, z) \quad \text{for any } z, x \text{ in } \boldsymbol{V}. \tag{44}$$

The property is true by assumption if both z and x are in $\boldsymbol{V}'$; if instead z and y are in $\boldsymbol{V} \setminus \boldsymbol{V}'$ we have

$$v(z) - v(x) \leq g(y) + S_a(y, z) - g(y) - S_a(y, x) \leq S_a(x, z),$$

where $y \in \boldsymbol{V}'$ is optimal for $v(x)$ and we have exploited the triangle inequality (35). If $z \notin \boldsymbol{V}'$ and $x \in \boldsymbol{V}'$, then (44) directly comes from the very definition of v. Finally, if $z \in \boldsymbol{V}'$ and $x \notin \boldsymbol{V}'$, we denote by y an optimal element in $\boldsymbol{V}'$ and use the triangle inequality to write

$$v(z) - v(x) = g(z) - g(y) - S_a(y, x) \leq S_a(y, z) - S_a(y, x) \leq S_a(x, z).$$

This concludes the proof of claim (44) and therefore shows, according to Proposition 6.8, that v is a subsolution in $\boldsymbol{V}$. Taking into account that $S_a(y, \cdot)$ is a solution in $\boldsymbol{V} \setminus \boldsymbol{V}'$, we also get, arguing as in Theorem 6.21, that v is a solution in $\boldsymbol{V} \setminus \boldsymbol{V}'$. Uniqueness follows from Proposition 6.6. ☐

7. Back to the network

In this section we switch our attention back to the network Γ, or in other terms, we give again visibility, besides the vertices, to the interior points of the arcs. We combine the global information gathered on the abstract graph with the outputs of the local analysis on the arcs of the network. We define an appropriate notion of Aubry set and provide a PDE characterization of its points.

Exploiting the richer (differentiable) structure of Γ, we establish, on the basis of our findings in the previous section, some regularity properties for critical subsolutions and solutions. This will generalize what is known for the continuous case in the framework of weak KAM theory; see for example [Fathi 2008]. Finally, we give specific uniqueness results and representation formulae for solutions on the network.

7A. ***Subsolutions and solutions on*** $\boldsymbol{\Gamma}$**.** The next result shows, as pointed out already in the Introduction, how the notion of solution to ($\mathcal{H}Ja$) can be recovered from the notion of subsolution. The relevance of the issue is that the latter just requires the usual subsolution property on any arc and continuity at the junctures. The argument significantly illustrates the interplay between the immersed network and underlying abstract graph.

Theorem 7.1. *Let $a \ge c$ and $y \in \Gamma$; then the maximal subsolution to* ($\mathcal{H}Ja$) *attaining a given value at y is a solution in* $\Gamma \setminus \{y\}$.

Proof. We can assume $y \in \Gamma \setminus \boldsymbol{V}$; otherwise the assertion is a consequence of Propositions 6.8 and 6.11 and Proposition 6.3 with $\boldsymbol{V}' = \{y\}$. It is not restrictive to take 0 as the value assigned at y. We therefore denote by v the maximal subsolution vanishing at y; see Proposition 3.12. We select $\gamma \in \mathcal{E}$ such that $y = \gamma(s_0)$ for some $s_0 \in (0, 1)$, and set $e = \Psi^{-1}(\gamma)$. We first assume that γ is not a closed arc. Since v must be in particular a subsolution in the arc γ, we have by Corollary 5.7

$$
\begin{aligned}
v(\gamma(1)) &\le \int_{s_0}^{1} \sigma_a^+(t)\,dt =: \beta, \\
v(\gamma(0)) &\le -\int_{0}^{s_0} \sigma_a^-(t)\,dt =: \alpha,
\end{aligned}
$$

where σ_a^+, σ_a^- are defined as in (8), (9). The maximal admissible trace g, in the sense of (41), on $\boldsymbol{V}' := \{\mathrm{o}(e), \mathrm{t}(e)\}$ dominated by α at $\mathrm{o}(e) = \gamma(0)$, and β at $\mathrm{t}(e) = \gamma(1)$, is

$$
\begin{aligned}
\alpha^* &:= \min\{\alpha,\ \beta + S_a(\mathrm{t}(e), \mathrm{o}(e))\}, \\
\beta^* &:= \min\{\beta,\ \alpha + S_a(\mathrm{o}(e), \mathrm{t}(e))\}.
\end{aligned}
$$

According to Proposition 6.8, Theorem 6.23 and Corollary 6.22, the function $w : \boldsymbol{V} \to \mathbb{R}$, defined as

$$
w(x) = \begin{cases} \alpha^* & \text{if } x = \mathrm{o}(e), \\ \beta^* & \text{if } x = \mathrm{t}(e) \\ \min\{\alpha^* + S_a(\mathrm{o}(e), x),\ \beta^* + S_a(\mathrm{t}(e), x)\} & \text{if } x \ne \mathrm{o}(e) \text{ and } x \ne \mathrm{t}(e), \end{cases}
$$

is the maximal subsolution to ($\mathcal{DFE}a$) on $\boldsymbol{V}$ agreeing with α^* and β^* at the vertices of e. It is in addition a solution in $\boldsymbol{V} \setminus \{\gamma(0), \gamma(1)\}$. By Proposition 6.3 it can thus be extended to a subsolution of ($\mathcal{H}Ja$)

in Γ, denoted by $\overline{w}$, which is in addition a solution in $\Gamma \setminus \{\gamma(0), \gamma(1)\}$. The function $\overline{w}$ is the maximal subsolution to $(\mathcal{H}Ja)$ taking the values α^* and β^* on the vertices of γ, but it does not necessarily vanish at y. We have in any case

$$v \le \overline{w} \quad \text{in } \Gamma. \tag{45}$$

To complete the proof, we need to suitably adjust $\overline{w}$ inside γ in order to attain the value 0 at y. To this end, we proceed by showing that the boundary data α^*, 0 and 0, β^* are admissible, in the sense of (17), for $(HJ_\gamma a)$ restricted to the subintervals $[0, s_0]$ and $[s_0, 1]$, respectively. In fact,

$$\alpha^* \le \alpha = -\int_0^{s_0} \sigma_a^-(t)\,dt, \tag{46}$$

and if a strict inequality prevails in the above formula, we get

$$\alpha^* = \int_{s_0}^1 \sigma_a^+(t)\,dt + S_a(\mathrm{t}(e), \mathrm{o}(e)). \tag{47}$$

Let us consider a cycle in $\boldsymbol{X}$ of the form $\xi \cup e$, where ξ is a path linking $\mathrm{t}(e)$ to $\mathrm{o}(e)$ with $\sigma_a(\xi) = S_a(\mathrm{t}(e), \mathrm{o}(e))$; see Corollary 6.10. Then $\sigma_a(\xi \cup e) \ge 0$ and consequently $S_a(\mathrm{t}(e), \mathrm{o}(e)) \ge -\sigma_a(e)$. By plugging this relation into (47) and recalling the definition of $\sigma_a(e)$, we get

$$\alpha^* \ge \int_{s_0}^1 \sigma_a^+(t)\,dt - \int_0^1 \sigma_a^+(t)\,dt = -\int_0^{s_0} \sigma_a^+(t)\,dt. \tag{48}$$

By combining (46) and (48) we have

$$\int_0^{s_0} \sigma_a^-(t)\,dt \le -\alpha^* \le \int_0^{s_0} \sigma_a^+(t)\,dt,$$

proving the claimed admissibility property in $[0, s_0]$. A straightforward modification of the previous argument shows the same in $[s_0, 1]$. Thus, there exists a function u on $\gamma([0, 1])$ uniquely determined by requiring $u \circ \gamma$ to be a solution to $(HJ_\gamma a)$ in $(0, s_0)$ and $(s_0, 1)$, and in addition taking the values α^*, 0 and β^* at $\gamma(0)$, y and $\gamma(1)$, respectively. This is also the maximal subsolution of $(HJ_\gamma a)$ in $(0, 1)$ taking such values at the boundary points and at $s = s_0$. The function

$$\overline{\overline{w}}(x) = \begin{cases} \overline{w} & \text{in } \Gamma \setminus \gamma([0,1]), \\ u & \text{in } \gamma([0,1]) \end{cases}$$

is a subsolution to $(\mathcal{H}Ja)$ in Γ and by the maximality property of u on γ and (45),

$$v \le \overline{\overline{w}} \quad \text{in } \Gamma,$$

which immediately implies $v = \overline{\overline{w}}$.

The function v is by construction a solution to $(\mathcal{H}Ja)$ in $\Gamma \setminus \{\gamma(0), y, \gamma(1)\}$. Moreover, taking into account Remark 5.2 and Proposition 5.6 applied to the subinterval $[0, s_0]$, we see that if $\overline{w}(\gamma(0)) = \alpha$ then $\overline{w}$ satisfies condition (iii) in the definition of solution to $(\mathcal{H}Ja)$ at $\gamma(0)$ with respect to the arc $\tilde{\gamma}$. If instead $\overline{w}(\mathrm{o}(e)) = \alpha + S_a(\mathrm{t}(e), \mathrm{o}(e))$ then again condition (iii) of the definition of solution is satisfied with respect to some arc different from γ, $\tilde{\gamma}$ because of Propositions 6.11 and 6.3. Similarly, we prove that v is a solution at $\gamma(1)$. This concludes the proof if γ is not a closed arc.

If instead γ is a closed arc, then we indicate by w the maximal periodic subsolution of ($HJ_\gamma a$) in (0, 1) vanishing at $s = s_0$; see Corollary 5.13. Arguing as in the first part of the proof, we see that the maximal subsolution v to ($\mathcal{H}Ja$) in Γ vanishing at y is given by

$$v(x) = \begin{cases} w(\gamma^{-1}(x)) & \text{in } \gamma([0, 1]), \\ w(\gamma(0)) + S_a(\gamma(0), x) & \text{in } \Gamma \setminus \gamma([0, 1]). \end{cases}$$

Taking into account the representation formulae for w provided in item (ii) of Corollary 5.13 and arguing again as in the first part of the proof, we show that v is a solution to ($\mathcal{H}Ja$) in $\Gamma \setminus \{y\}$, as was claimed. $\square$

7B. *Aubry set in* Γ. We define the Aubry set $\mathcal{A}_\Gamma$ on the network as

$$\mathcal{A}_\Gamma := \{x \in \mathbb{R}^N \mid x = \Psi(e)(t) \text{ for some } e \in \mathcal{A}^*_{\boldsymbol{X}}, t \in [0, 1]\}. \tag{49}$$

One could also consider a lift of $\mathcal{A}_\Gamma$ to the tangent bundle $T\Gamma$, as in the continuous case. For example, this could be useful to study the analogues in this setting of Mather's measures, Mather sets, minimal average actions, etc. (see for example [Fathi 2008; Sorrentino 2015] for precise definitions); this discussion, however, would go beyond our current objectives, so we decided to postpone it to a future investigation.

Remark 7.2. We point out for later use that the support of an arc γ belongs to $\mathcal{A}_\Gamma$ if and only if $\gamma = \Psi(e)$ and at least one between e or $-e$ is in $\mathcal{A}^*_{\boldsymbol{X}}$.

The first lemma regards subsolutions to the critical equation on $\boldsymbol{X}$. Briefly, it says that — analogously to what happens in the continuous case, see [Fathi 2008] — the differential of a critical subsolution is prescribed on the Aubry set and that critical subsolutions are never strict on the Aubry set. On the other hand, it is always possible to find critical subsolutions that are strict outside the Aubry set. This will be used in the next subsection to obtain the same results on networks. See Theorems 7.5 and 7.6.

Lemma 7.3. *Given a subsolution u to ($\mathcal{D}FEc$), one has*

$$\langle \mathrm{d}u, e\rangle = \sigma_c(e) \quad \textit{for any } e \in \mathcal{A}^*_{\boldsymbol{X}}. \tag{50}$$

Furthermore, there exists a subsolution w to ($\mathcal{D}FEc$) with

$$\langle \mathrm{d}w, e\rangle < \sigma_c(e) \quad \textit{for any } e \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}. \tag{51}$$

Proof. Let u be a critical subsolution and assume for purposes of contradiction that

$$\langle \mathrm{d}u, \bar{e}\rangle < \sigma_c(\bar{e}) \quad \text{for some } \bar{e} \in \mathcal{A}^*_{\boldsymbol{X}}.$$

By the very definition of Aubry set, we can find a cycle $\xi = (e_i)_{i=1}^M$ such that $\bar{e} = e_j$ for some $j = 1, \dots, M$ and $\sigma_c(\xi) = 0$. Taking into account that u is a subsolution, we have

$$\langle \mathrm{d}u, e_i\rangle \le \sigma_c(e_i) \quad \text{for } i \ne j \qquad \text{and} \qquad \langle \mathrm{d}u, e_j\rangle < \sigma_c(e_j).$$

This implies

$$0 = \sum_i \langle \mathrm{d}u, e_i \rangle < \sum_i \sigma_c(e_i) = \sigma_c(\xi) = 0,$$

which is impossible. We pass to the second part of the statement. We start constructing for any $e_0 \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}$ a critical subsolution u_{e_0} with

$$\langle\, du_{e_0}, e_0 \rangle < \sigma_c(e_0). \tag{52}$$

The argument will be organized taking into account the classification of edges in $\mathcal{A}^*_{\boldsymbol{X}}$ provided in Remark 6.19. If $\mathrm{t}(e_0) \notin \mathcal{A}_{\boldsymbol{X}}$, then we set $u_{e_0} = S_c(\mathrm{t}(e_0), \cdot\,)$; according to Lemma 6.20, u_{e_0} is not a critical solution at $\mathrm{t}(e_0)$ which implies (52). If $\mathrm{t}(e_0) \in \mathcal{A}_{\boldsymbol{X}}$, we consider the critical subsolutions $S_c(\mathrm{t}(e_0), \cdot\,)$ and $-S_c(\,\cdot\,, \mathrm{t}(e_0))$; see Proposition 6.11 and Corollary 6.12. Taking into account the characterization of $\mathcal{A}_{\boldsymbol{X}}$ given in Lemma 6.20, we have

$$\begin{aligned} -S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) &= \quad S_c(\mathrm{t}(e_0), \mathrm{t}(e_0)) - S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) \quad \le \sigma_c(e_0), \\ S_c(\mathrm{o}(e_0), \mathrm{t}(e_0)) &= -S_c(\mathrm{t}(e_0), \mathrm{t}(e_0)) + S_c(\mathrm{o}(e_0), \mathrm{t}(e_0)) \le \sigma_c(e_0). \end{aligned}$$

If equality prevails in both above formulae, we get

$$S_c(\mathrm{o}(e_0), \mathrm{t}(e_0)) + S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) = 0,$$

which is possible if and only if both $\mathrm{o}(e_0)$ and $\mathrm{t}(e_0)$ are in the Aubry set and belong to the same static class. If this is not the case, we satisfy (52) up to choosing u_{e_0} equal to $S_c(\mathrm{t}(e_0), \cdot\,)$ or $-S_c(\,\cdot\,, \mathrm{t}(e_0))$. If instead the two vertices are in the same static class, we claim that

$$S_c(\mathrm{t}(e_0), \mathrm{t}(e_0)) - S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) = -S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) < \sigma_c(e_0). \tag{53}$$

In fact, we know, by the very definition of static class, that there is a path ξ linking $\mathrm{t}(e_0)$ to $\mathrm{o}(e_0)$ with all the edges belonging to $\mathcal{A}^*_{\boldsymbol{X}}$. Therefore, using Lemma 6.20 and the first part of the statement that we have just proven, applied to the critical subsolution $-S_c(\,\cdot\,, \mathrm{o}(e_0))$, we have that

$$S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) = -S_c(\mathrm{o}(e_0), \mathrm{o}(e_0)) + S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) = \sigma_c(\xi).$$

Were (53) false, we should further have

$$0 = -S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) + S_c(\mathrm{t}(e_0), \mathrm{o}(e_0)) = \sigma_c(\xi \cup e_0)$$

and consequently $e_0 \in \mathcal{A}^*_{\boldsymbol{X}}$, which is impossible. Formula (52) is therefore satisfied with $u_{e_0} = S_c(\mathrm{t}(e_0), \cdot\,)$. This completes the proof of (52).

We conclude arguing along the same lines as Theorem 6.16. Given $e \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}$, we denote by u_e a critical subsolution satisfying (52) with e in place of e_0. We choose positive constants λ_e for $e \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}$, summing to 1, and define a critical subsolution via

$$w = \sum_{e \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}} \lambda_e\, u_e.$$

Given $e_0 \in \boldsymbol{E} \setminus \mathcal{A}^*_{\boldsymbol{X}}$, we have

$$\langle \mathrm{d}w, e_0\rangle = \sum_{e \neq e_0} \lambda_e \, \langle \mathrm{d}u_e, e_0\rangle + \lambda_{e_0} \, \langle \mathrm{d}u_{e_0}, e_0\rangle < \sigma_c(e_0),$$

as we wished to prove. □

We derive a PDE characterization of points in the Aubry set, generalizing a property of the continuous case.

Proposition 7.4. *The maximal subsolution to* ($\mathcal{H}Jc$) *taking a given value at a point* $y \in \Gamma$ *is a critical solution on the whole network if and only if* $y \in \mathcal{A}_\Gamma$.

Proof. If $y \in \boldsymbol{V}$, the assertion comes from Lemma 6.20; we can then assume from now on that $y \in \Gamma \setminus \boldsymbol{V}$. We prescribe, without loss of generality, the value 0 at y, and denote by v the maximal subsolution vanishing at y; see Proposition 3.12. We denote by γ an arc whose support contains y.

We first assume that γ is not a closed curve. Taking into account Theorem 7.1, it is enough to show that v is a solution at y if and only if $y \in \mathcal{A}_\Gamma$. Looking at the proof of Theorem 7.1 (we adopt the same notations), we see that the solution property at y is in turn equivalent to the following: the solution of ($HJ_\gamma c$) in (0, 1) taking the values $v(\gamma(0))$, $v(\gamma(1))$ at 0, 1, respectively, vanishes at $s = s_0$. In light of Proposition 5.5, this boils down to showing

$$\min\{v(\gamma(0)) + \boldsymbol{A},\ v(\gamma(1)) - \boldsymbol{B}\} = 0, \tag{54}$$

where σ_c^+, σ_c^- are defined as in (8), (9), respectively, and

$$\boldsymbol{A} = \int_0^{s_0} \sigma_c^+(t)\, dt, \quad \boldsymbol{B} = \int_{s_0}^1 \sigma_c^-(t)\, dt.$$

Taking into account the proof of Theorem 7.1, we know that

$$v(\gamma(0)) = \min\{-\boldsymbol{D},\ \boldsymbol{C} + S_c(\gamma(1), \gamma(0))\}, \tag{55}$$

$$v(\gamma(1)) = \min\{\boldsymbol{C},\ -\boldsymbol{D} + S_c(\gamma(0), \gamma(1))\}, \tag{56}$$

where

$$\boldsymbol{C} = \int_{s_0}^1 \sigma_c^+(t)\, dt, \quad \boldsymbol{D} = \int_0^{s_0} \sigma_c^-(t)\, dt.$$

Then

$$v(\gamma(0)) + \boldsymbol{A} = \begin{cases} \int_0^{s_0} [\sigma_c^+(t) - \sigma_c^-(t)]\, dt & \text{if } v(\gamma(0)) = -\boldsymbol{D}, \\ \int_0^1 \sigma_c^+(t)\, dt + S_c(\gamma(1), \gamma(0)) & \text{if } v(\gamma(0)) = \boldsymbol{C} + S_c(\gamma(1), \gamma(0)) \end{cases} \tag{57}$$

and

$$v(\gamma(1)) - \boldsymbol{B} = \begin{cases} \int_{s_0}^1 [\sigma_c^+(t) - \sigma_c^-(t)]\, dt & \text{if } v(\gamma(1)) = \boldsymbol{C}, \\ -\int_0^1 \sigma_c^-(t)\, dt + S_c(\gamma(0), \gamma(1)) & \text{if } v(\gamma(1)) = -\boldsymbol{D} + S_c(\gamma(0), \gamma(1)). \end{cases} \tag{58}$$

Exploiting the property that $\sigma_c(\xi) \geq 0$ for any cycle ξ in $\boldsymbol{X}$, we see that

$$S_c(\gamma(0), \gamma(1)) \geq -\sigma_c(-e) = \int_0^1 \sigma_c^-(t)\, dt,$$

$$S_c(\gamma(1), \gamma(0)) \geq -\sigma_c(e) = -\int_0^1 \sigma_c^+(t)\, dt.$$

Equality holds in the first formula if and only if there is a cycle ξ with $-e \subset \xi$ and $\sigma_c(\xi) = 0$, and in the second one if and only if there a cycle η with $e \subset \xi$ and $\sigma_c(\eta) = 0$. We in addition have that

$$\int_0^{s_0} [\sigma_c^+(t) - \sigma_c^-(t)]\, dt = 0 \quad \text{or} \quad \int_{s_0}^1 [\sigma_c^+(t) - \sigma_c^-(t)]\, dt = 0$$

if and only if $c = a_\gamma$, and in this case both e and $-e$ belong to $\mathcal{A}^*_{\boldsymbol{X}}$. In light of the above remarks, (57) and (58), we conclude that (54) holds if and only if $y \in \mathcal{A}_\Gamma$.

This concludes the proof when γ is not a closed arc. The argument for γ a closed arc goes along the same lines just adapting the representation formulae for solutions of $(HJ_\gamma c)$ and taking into account Corollary 5.13. □

7C. *Regularity results for critical subsolutions.* We state and prove the main regularity results of this section. They can be considered as a generalization to the network setting of the results in [Fathi and Siconolfi 2004].

Theorem 7.5. *Any critical subsolution $u : \Gamma \to \mathbb{R}$ is of class C^1 in $\mathcal{A}_\Gamma \setminus \boldsymbol{V}$, and all such subsolutions possess the same differential in $\mathcal{A}_\Gamma \setminus \boldsymbol{V}$.*

Proof. Let u be a critical subsolution on Γ and $\gamma = \Psi(e)$ an arc with $e \in \mathcal{A}^*_{\boldsymbol{X}}$. According to Lemma 7.3 and formula (50),

$$u(\gamma(1)) - u(\gamma(0)) = \sigma_c(e).$$

Therefore $u \circ \gamma$ is the maximal subsolution taking the value $u(\gamma(0))$ at $s = 0$ and, according to Proposition 5.6, has the form

$$u(\gamma(s)) = \int_0^s \sigma_c^+(t)\, dt,$$

where σ_c^+ is as in (8) with H_γ in place of H and c in place of a. We deduce that $s \mapsto u(\gamma(s))$ is of class C^1 for $t \in (0, 1)$ and for any $x = \gamma(t_0)$, with $t_0 \in (0, 1)$, the differential $D_\Gamma u(x)$ is uniquely determined among the elements of $T^*_\Gamma(x)$ by the condition

$$(D_\Gamma u(x), \dot\gamma(t_0)) = \frac{d}{dt} u(\gamma(t))|_{t=t_0} = \sigma_c^+(t_0). \qquad \square$$

Moreover:

Theorem 7.6. *For any critical subsolution w on $\boldsymbol{X}$, there exists a critical subsolution u on Γ, with $w = u$ on $\boldsymbol{V}$, which is of class C^1 in $\Gamma \setminus \boldsymbol{V}$. There exists in addition a critical subsolution v on Γ of class $C^1(\Gamma \setminus \boldsymbol{V})$ satisfying*

$$\mathcal{H}(x, D_\Gamma v(x)) < c \quad \textit{for } x \in \Gamma \setminus (\mathcal{A}_\Gamma \cup \boldsymbol{V}).$$

Proof. Let w be a critical subsolution in $\boldsymbol{X}$. Given any arc $\gamma = \Psi(e)$, we know, see Proposition 6.2, that $w(\gamma(0))$ and $w(\gamma(1))$ satisfy the compatibility condition (17), so that

$$w(\gamma(0)) + \int_0^1 \sigma_c^-(t)\,dt \le w(\gamma(1)) \le w(\gamma(0)) + \int_0^1 \sigma_c^+(t)\,dt, \tag{59}$$

where σ_c^+, σ_c^- are defined as in (8), (9) with H_γ, c in place of H, a, respectively. We can therefore find $\lambda \in [0, 1]$ with

$$w(\gamma(1)) = w(\gamma(0)) + \int_0^1 [\lambda\,\sigma_c^-(t) + (1-\lambda)\,\sigma_c^+(t)]\,dt, \tag{60}$$

and the function

$$s \mapsto w(\gamma(0)) + \int_0^s [\lambda\,\sigma_c^-(t) + (1-\lambda)\,\sigma_c^+(t)]\,dt \tag{61}$$

is a subsolution of class C^1 to $H_\gamma = c$ in $(0, 1)$ taking the values $w(\gamma(0))$ and $w(\gamma(1))$ at $s = 0$ and $s = 1$, respectively. This shows the first part of the assertion.

As far as the second claim is concerned, we proceed by taking a critical subsolution w satisfying (51). This implies that strict inequalities prevail in formula (59) whenever $\gamma = \Psi(e)$ with e, $-e$ not in $\mathcal{A}^*_{\boldsymbol{X}}$. The λ appearing in (60) can be consequently taken in $(0, 1)$, so that the function defined in (61) is a strict subsolution to $H_\gamma = c$. This concludes the proof in light of Remark 7.2. □

Remark 7.7. Notice that if we apply the procedure of the first part of the previous result starting with a critical solution rather than a critical subsolution, then the property of being a solution could be possibly false for the regularized function.

7D. ***Representation formulae and uniqueness results on the network.*** In this section, we want to provide representation formulae and uniqueness results with traces that are not necessarily defined on vertices, but on a general subset of the network Γ. To this aim, we extend S_a, for $a \ge c$, from $\boldsymbol{V}$ to the whole of Γ defining a semidistance intrinsically related to $\mathcal{H}$ and the level a. This is basically the same object introduced in [Schieborn and Camilli 2013]. We do not develop here any further the metric point of view, but just use it to establish an admissibility condition for data assigned on subsets of Γ, and provide representation formulae.

Given a portion of arc $\gamma|_{[s_1,s_2]}$, for $0 \le s_1 \le s_2 \le 1$, we define

$$\ell_a(\gamma|_{[s_1,s_2]}) = \int_{s_1}^{s_2} (\sigma_a^+)^\gamma(t)\,dt,$$

where $(\sigma_a^+)^\gamma$ is defined as in (8). We get in particular, for the whole arc, the relation

$$\ell_a(\gamma) = \sigma_a(\Psi^{-1}(\gamma)) \quad \text{for any } \gamma \in \mathcal{E}. \tag{62}$$

We define ℓ_a for a curve on Γ given by a finite number of concatenated arcs or portions of arcs as the sum of the lengths of the arcs or portion of arcs making it up. We introduce the related geodesic (semi-)distance on Γ via

$$S_a^\Gamma(x, y) = \min\{\ell_a(\xi) \mid \xi \text{ a union of concatenated arcs linking } x \text{ to } y\}. \tag{63}$$

We deduce from the results on σ_a and (62) the following lemma.

Lemma 7.8. (i) *If $x \neq y$ are in $\boldsymbol{V}$, then $S_a(x, y) = S_a^\Gamma(x, y)$.*

(ii) *If ξ is a closed curve on Γ, then $\ell_a(\xi) \geq 0$.*

It is easy to check that the maximal subsolution v to ($\mathcal{D}FEa$) vanishing at $y \in \Gamma$ given in Theorem 7.1 and Proposition 7.4 is

$$v(x) = S_a^\Gamma(y, x) \quad \text{for any } a \geq c,\ \ x \in \Gamma.$$

We derive, taking also into account Proposition 6.8, that for a continuous function $u : \Gamma \to \mathbb{R}$, the condition

$$u(x) - u(y) \leq S_a^\Gamma(y, x) \quad \text{for any pair } x,\, y \text{ in } \Gamma' \tag{64}$$

is necessary and sufficient for being a subsolution to ($\mathcal{H}Ja$). Given a function g defined on a subset Γ' of Γ, we therefore introduce the following admissibility condition for ($\mathcal{D}FEa$):

$$g(x) - g(y) \leq S_a^\Gamma(x, y) \quad \text{for any } x,\, y \text{ in } \Gamma'. \tag{65}$$

We give in the next theorem a couple of examples of uniqueness results for solutions to ($\mathcal{D}FEa$), and corresponding representation formulae, one can obtain prescribing values on subsets not necessarily contained in $\boldsymbol{V}$. Further results are reachable along the same lines. Similar formulae, even if for subsets of vertices and just in the supercritical case, have been already obtained in [Schieborn and Camilli 2013].

Theorem 7.9. *Let Γ' be a closed subset of Γ and g an admissible trace defined on it, in the sense of* (65). *We set*

$$v(x) = \min\{g(y) + S_a^\Gamma(y, x) \mid y \in \Gamma'\}.$$

(i) *Critical case: if $a = c$ and $\Gamma' \subset \mathcal{A}_\Gamma$ with*

$$\Gamma' \cap \gamma([0, 1]) \neq \varnothing \quad \textit{for any } \gamma \textit{ with } \Psi^{-1}(\gamma) \in \mathcal{A}^*_{\boldsymbol{X}}, \tag{66}$$

then v is the unique solution in Γ to $\mathcal{H}(x, Du) = c$ agreeing with g on Γ'.

(ii) *Supercritical case: if $a > c$, then v is uniquely characterized by the properties of being in $C(\Gamma, \mathbb{R})$, being a solution of ($\mathcal{H}Ja$) in $\Gamma \setminus \Gamma'$, and agreeing with g on Γ'.*

Proof. The solution property of v in both cases, in Γ and $\Gamma \setminus \Gamma'$ respectively, follows directly from being a subsolution in Γ, in light of (64), and satisfying the subtangent test as a minimum of solutions, in Γ and $\Gamma \setminus \Gamma'$ respectively. In addition v is the maximal solution (in Γ or $\Gamma \setminus \Gamma'$) agreeing with g on Γ' in light of Theorem 7.1, Proposition 7.4, and the admissibility condition (65).

Now, assume u to be another solution taking the value g on Γ'; by adapting the backward procedure explained in Proposition 6.15 and Theorem 6.21, we construct, for any $x \in \Gamma \setminus \Gamma'$, a curve ξ made up by concatenated arcs or portion of arcs starting at some point $y \in \Gamma'$ and ending at x with

$$u(x) = g(y) + \ell_a(\xi) \geq v(x).$$

In the critical case, condition (66) plays a crucial role for this. The maximality property of v then implies that equality must hold in the above formula. ☐

Appendix

Proof of Proposition 3.11. Taking into account that for any $\gamma \in \mathcal{E}$ (which is a finite set) $w \circ \gamma$ is Lipschitz-continuous in [0, 1], thanks to the coercivity condition (Hγ2), we deduce that there exists $L > 0$ such that, for any given subsolution w,

$$|w(\gamma(s_2)) - w(\gamma(s_1))| \le L\, \ell(\gamma|_{[s_1,s_2]}) \quad \text{for all } \gamma \in \mathcal{E}, \text{ and } s_1 \le s_2 \in [0,1]; \tag{67}$$

hereafter ℓ indicates the Euclidean length of curves in $\mathbb{R}^N$.

We proceed by considering x and y in Γ and a finite sequence of concatenated arcs $\gamma_1, \dots, \gamma_M$, for some index M, that realize the geodesic distance $d_\Gamma(x, y)$. More specifically, we assume that $x = \gamma_1(t_x)$, $y = \gamma_M(t_y)$ with t_x, t_y in [0, 1] and that

$$d_\Gamma(x, y) \;=\; \ell(\gamma_1|_{[t_x,1]}) + \sum_{i=2}^{M-1} \ell(\gamma_i) + \ell(\gamma_M|_{[0,t_y]}).$$

In the remainder of the proof we assume that $M > 2$ in order to ease the notation (the other cases can be treated analogously).

We deduce from (67) that

$$\begin{aligned} |w(y) - w(x)| &\le |w(\gamma_1(1)) - w_1(\gamma_1(t_x))| + \sum_{i=2}^{M-1} |w(\gamma_i(1)) - w(\gamma_i(0))| + |w(\gamma_M(t_y)) - w_1(\gamma_M(0))| \\ &\le L\left[\ell(\gamma_1|_{[t_x,1]}) + \sum_{i=2}^{M-1} \ell(\gamma_i) + \ell(\gamma_M|_{[0,t_y]})\right] = L\, d_\Gamma(x, y). \end{aligned}$$

□

Proof of Proposition 5.5. We denote by w the function appearing in the statement. If $a = a_\gamma$, the assertion comes from (11) and Proposition 5.3. Instead, if $a > a_\gamma$, the function w is an a.e. subsolution, being the minimum of two C^1 (sub-)solutions. Using a basic property in viscosity solutions theory, it is also a supersolution, as a minimum of supersolutions. Moreover, $w(0) = \alpha$ and $w(1) = \beta$ hold thanks to (17).

Finally, the function $s \mapsto \int_0^s \sigma^+_{a_\gamma}$ is a strict subsolution to $(HJ_\gamma a)$, and this implies by an argument going back to [Ishii 1987] that the Dirichlet problem with admissible data α, β is uniquely solved. □

Proof of Proposition 5.6. If $a = a_\gamma$, then, as already pointed out in Proposition 5.3, the solution is unique up to additive constants; hence it is automatically given by (20) once the value $w(0)$ is assigned.

Therefore, from now on we can assume that $a > a_\gamma$. By Proposition 5.5,

$$w(s) = \min\left\{ w(0) + \int_0^s \sigma^+_a(t)\, dt,\; w(1) - \int_s^1 \sigma^-_a(t)\, dt \right\} \quad \text{for any } s.$$

We claim that if

$$w(s_0) = w(1) - \int_{s_0}^1 \sigma^-_a(t)\, dt \tag{68}$$

for some $s_0 \in (0, 1)$, then

$$w(s) = w(1) - \int_s^1 \sigma^-_a(t)\, dt \quad \text{for any } s \in (s_0, 1].$$

Assume by contradiction that there exists $s_1 > s_0$ such that

$$w(0)+\int_0^{s_1}\sigma_a^+(t)\,dt = w(0)+\int_0^{s_0}\sigma_a^+(t)\,dt+\int_{s_0}^{s_1}\sigma_a^+(t)\,dt < w(1)-\int_{s_1}^1\sigma^-(t)\,dt;$$

this implies

$$w(0)+\int_0^{s_0}\sigma_a^+(t)\,dt < w(1)-\int_{s_1}^1\sigma_a^-(t)\,dt-\int_{s_0}^{s_1}\sigma_a^+(t)\,dt. \tag{69}$$

It is apparent that

$$\int_{s_0}^{s_1}\sigma_a^+(t)\,dt > \int_{s_0}^{s_1}\sigma_a^-(t)\,dt$$

and we can consequently deduce from (69) that

$$w(0)+\int_0^{s_0}\sigma_a^+(t)\,dt < w(1)-\int_{s_1}^1\sigma_a^-(t)\,dt-\int_{s_0}^{s_1}\sigma_a^-(t)\,dt = w(1)-\int_{s_0}^1\sigma^-(t)\,dt,$$

in contrast with (68). We assume, for purposes of contradiction, that (68) holds true for some $s_0 \in (0,1)$. Since $a > a_\gamma$, we can take p_0 with $H(1,p_0) < a$. If w is not of the form (20), then, owing to the previous claim, we can fix s_0 in such a way that

$$w(s) = w(1)-\int_s^1\sigma_a^-(t)\,dt \quad\text{and}\quad H(s,p_0) < a$$

for $s \in [s_0,1]$. This implies

$$\varphi(s) := w(1)+p_0(s-1) \le w(1)-\int_s^1\sigma_a^-(t)\,dt = w(s)$$

for $s \in [s_0,1]$, and consequently φ is a constrained subtangent to w at 1 with

$$H(1,\varphi'(1)) = H(1,p_0) < 1,$$

contradicting (19). We deduce that w is of the form (20) showing the first part of the assertion.

Conversely, if w is of the form (20), then it is of class C^1 in $(0,1)$ with $w'(s)=\sigma_a^+(s)$. Consequently, any constrained subtangent φ at $t=1$ must satisfy

$$w(1)-\int_s^1\varphi'\,dt = \varphi(s) \le w(s) = w(1)-\int_s^1\sigma_a^+\,dt$$

for s sufficiently close to 1. This implies

$$\int_s^1\varphi'\,dt \ge \int_s^1\sigma_a^+\,dt$$

and shows the existence of a sequence s_n contained in $(0,1)$ and converging to 1 as n goes to infinity, with $\varphi'(s_n) \ge \sigma_a^+(s_n)$. Passing to the limit as n goes to infinity, we get $\varphi'(1) \ge \sigma_a^+(1)$. We deduce from this the inequality (19) and conclude the proof. □

Proof of Proposition 5.11. If $a = c_\gamma = a_\gamma$ then the integrals in (21) coincide in light of (11); then they must both vanish, and this shows the assertion. Assume now that $c_\gamma > a_\gamma$ and also assume for purposes of contradiction that strict inequalities prevail instead in (21). Then, we can find $\lambda \in (0, 1)$ with

$$\int_0^1 [\lambda\, \sigma^+_{c_\gamma}(t) + (1-\lambda)\, \sigma^-_{c_\gamma}(t)]\, dt = 0.$$

Taking into account that $\sigma^+_{c_\gamma}(t) > \sigma^-_{c_\gamma}(t)$ for any t, this implies

$$s \mapsto \int_0^s [\lambda\, \sigma^+_{c_\gamma}(t) + (1-\lambda)\, \sigma^-_{c_\gamma}(t)]\, dt$$

is a strict periodic subsolution to $H = c_\gamma$. This is impossible by the very definition of c_γ. □

Proof of Corollary 5.13. The unique point to check is that the values $\alpha + \beta$ at $s = 0$ and α at $s = s_0$ are admissible, in the sense of (17), for $(HJ_\gamma a)$ in $(0, s_0)$, and the same holds true in $(s_0, 1)$ for the values α at $s = s_0$ and $\alpha + \beta$ at $s = 1$. The argument is the same for the two subintervals. We therefore focus on $(s_0, 1)$.

If $u(1) - u(s_0) = \beta = \int_{s_0}^1 \sigma^+_a(t)\, dt$, the compatibility property is immediate and the solution in $(s_0, 1)$ is given by (23), as asserted in item (ii) of the statement. Let us instead assume

$$u(1) - u(s_0) = \beta = -\int_0^{s_0} \sigma^-_a(t)\, dt < \int_{s_0}^1 \sigma^+_a(t)\, dt. \tag{70}$$

We have by Lemma 5.9, $\int_0^1 \sigma^-_a(t)\, dt \le 0$ and consequently

$$u(1) - u(s_0) \ge \int_{s_0}^1 \sigma^-_a(t)\, dt.$$

The last inequality plus (70) shows the claimed admissibility property. □

Proof of Proposition 6.1. Let w be a solution to $(\mathcal{H}Ja)$ with trace u on $\boldsymbol{V}$. By the very definition of solution, given any arc γ, we know $w \circ \gamma$ is a solution to $H_\gamma = a$ in $(0, 1)$ taking the values $u(\gamma(0))$ and $u(\gamma(1))$ at 0 and 1, respectively. This implies that such boundary values are admissible with respect to H_γ, in the sense of formula (17) with H_γ in place of H. By the uniqueness property showcased in Proposition 5.5, the values of w on the support of γ are therefore uniquely determined by $u(\gamma(0))$, $u(\gamma(1))$ and H_γ. Since the arc γ has been arbitrarily chosen, we can hence conclude the asserted uniqueness. □

Acknowledgments

This work has been supported by the INdAM-GNAMPA research project "Fenomeni asintotici e omogeneizzazione". Siconolfi acknowledges the Progetto Ateneo 2015, Rome *La Sapienza* University: Asintotica e omogeneizzazione di dinamiche Hamiltoniane". Sorrentino acknowledges the PRIN-2012-74FYK7 grant "Variational and perturbative aspects of nonlinear differential problems". The authors are grateful to the anonymous referees for their valuable comments and suggestions.

References

[Achdou et al. 2013] Y. Achdou, F. Camilli, A. Cutrì, and N. Tchou, "Hamilton–Jacobi equations constrained on networks", *NoDEA Nonlinear Differential Equations Appl.* **20**:3 (2013), 413–445. MR Zbl

[Bardi and Capuzzo-Dolcetta 1997] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*, Birkhäuser, Boston, 1997. MR Zbl

[Barles 1994] G. Barles, *Solutions de viscosité des équations de Hamilton–Jacobi*, Mathématiques & Applications **17**, Springer, 1994. MR Zbl

[Barles et al. 2013] G. Barles, A. Briani, and E. Chasseigne, "A Bellman approach for two-domains optimal control problems in $\mathbb{R}^N$", *ESAIM Control Optim. Calc. Var.* **19**:3 (2013), 710–739. MR Zbl

[Barles et al. 2014] G. Barles, A. Briani, and E. Chasseigne, "A Bellman approach for regional optimal control problems in $\mathbb{R}^N$", *SIAM J. Control Optim.* **52**:3 (2014), 1712–1744. MR Zbl

[Bernard and Buffoni 2006] P. Bernard and B. Buffoni, "The Monge problem for supercritical Mañé potentials on compact manifolds", *Adv. Math.* **207**:2 (2006), 691–706. MR Zbl

[Bernard and Buffoni 2007] P. Bernard and B. Buffoni, "Optimal mass transportation and Mather theory", *J. Eur. Math. Soc. (JEMS)* **9**:1 (2007), 85–121. MR Zbl

[Bressan and Hong 2007] A. Bressan and Y. Hong, "Optimal control problems on stratified domains", *Netw. Heterog. Media* **2**:2 (2007), 313–331. MR Zbl

[Camilli and Marchi 2013] F. Camilli and C. Marchi, "A comparison among various notions of viscosity solution for Hamilton–Jacobi equations on networks", *J. Math. Anal. Appl.* **407**:1 (2013), 112–118. MR Zbl

[Camilli et al. 2013] F. Camilli, D. Schieborn, and C. Marchi, "Eikonal equations on ramified spaces", *Interfaces Free Bound.* **15**:1 (2013), 121–140. MR Zbl

[Davini et al. 2016] A. Davini, A. Fathi, R. Iturriaga, and M. Zavidovique, "Convergence of the solutions of the discounted equation: the discrete case", *Math. Z.* **284**:3-4 (2016), 1021–1034. MR Zbl

[Fathi 2008] A. Fathi, "Weak KAM Theorem in Lagrangian dynamics", lecture notes, 2008, https://www.math.u-bordeaux.fr/~pthieull/Recherche/KamFaible/Publications/Fathi2008_01.pdf.

[Fathi and Siconolfi 2004] A. Fathi and A. Siconolfi, "Existence of C^1 critical subsolutions of the Hamilton–Jacobi equation", *Invent. Math.* **155**:2 (2004), 363–388. MR Zbl

[Fathi and Siconolfi 2005] A. Fathi and A. Siconolfi, "PDE aspects of Aubry–Mather theory for quasiconvex Hamiltonians", *Calc. Var. Partial Differential Equations* **22**:2 (2005), 185–228. MR Zbl

[Galise et al. 2015] G. Galise, C. Imbert, and R. Monneau, "A junction condition by specified homogenization and application to traffic lights", *Anal. PDE* **8**:8 (2015), 1891–1929. MR Zbl

[Garavello and Piccoli 2006] M. Garavello and B. Piccoli, *Traffic flow on networks: conservation laws models*, AIMS Series on Applied Mathematics **1**, American Institute of Mathematical Sciences, Springfield, MO, 2006. MR Zbl

[Gomes 2005] D. A. Gomes, "Viscosity solution methods and the discrete Aubry–Mather problem", *Discrete Contin. Dyn. Syst.* **13**:1 (2005), 103–116. MR Zbl

[Imbert and Monneau 2016] C. Imbert and R. Monneau, "Quasi-convex Hamilton–Jacobi equations posed on junctions: the multi-dimensional case", preprint, 2016. arXiv

[Imbert and Monneau 2017] C. Imbert and R. Monneau, "Flux-limited solutions for quasi-convex Hamilton–Jacobi equations on networks", *Ann. Sci. Éc. Norm. Supér.* (4) **50**:2 (2017), 357–448. MR Zbl

[Imbert et al. 2013] C. Imbert, R. Monneau, and H. Zidani, "A Hamilton–Jacobi approach to junction problems and application to traffic flows", *ESAIM Control Optim. Calc. Var.* **19**:1 (2013), 129–166. MR Zbl

[Ishii 1987] H. Ishii, "A simple, direct proof of uniqueness for solutions of the Hamilton–Jacobi equations of eikonal type", *Proc. Amer. Math. Soc.* **100**:2 (1987), 247–251. MR Zbl

[Lions and Souganidis 2016] P.-L. Lions and P. Souganidis, "Viscosity solutions for junctions: well posedness and stability", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **27**:4 (2016), 535–545. MR Zbl

[Nash 1956] J. Nash, "The imbedding problem for Riemannian manifolds", *Ann. of Math.* (2) **63** (1956), 20–63. MR Zbl

[Pokornyi and Borovskikh 2004] Y. V. Pokornyi and A. V. Borovskikh, "Differential equations on networks (geometric graphs)", *J. Math. Sci.* (*N. Y.*) **119**:6 (2004), 691–718. MR Zbl

[Rao et al. 2014] Z. Rao, A. Siconolfi, and H. Zidani, "Transmission conditions on interfaces for Hamilton–Jacobi–Bellman equations", *J. Differential Equations* **257**:11 (2014), 3978–4014. MR Zbl

[Schieborn and Camilli 2013] D. Schieborn and F. Camilli, "Viscosity solutions of Eikonal equations on topological networks", *Calc. Var. Partial Differential Equations* **46**:3-4 (2013), 671–686. MR Zbl

[Soner 1986] H. M. Soner, "Optimal control with state-space constraint, I", *SIAM J. Control Optim.* **24**:3 (1986), 552–561. MR Zbl

[Sorrentino 2015] A. Sorrentino, *Action-minimizing methods in Hamiltonian dynamics: an introduction to Aubry–Mather theory*, Mathematical Notes **50**, Princeton University Press, 2015. MR Zbl

[Su and Thieullen 2015] X. Su and P. Thieullen, "Convergence of discrete Aubry–Mather model in the continuous limit", preprint, 2015. arXiv

[Sunada 2013] T. Sunada, *Topological crystallography: with a view towards discrete geometric analysis*, Surveys and Tutorials in the Applied Mathematical Sciences **6**, Springer, 2013. MR Zbl

[Zavidovique 2010] M. Zavidovique, "Existence of $C^{1,1}$ critical subsolutions in discrete weak KAM theory", *J. Mod. Dyn.* **4**:4 (2010), 693–714. MR Zbl

[Zavidovique 2012] M. Zavidovique, "Strict sub-solutions and Mañé potential in discrete weak KAM theory", *Comment. Math. Helv.* **87**:1 (2012), 1–39. MR Zbl

Received 5 Dec 2016. Revised 25 May 2017. Accepted 10 Aug 2017.

ANTONIO SICONOLFI: siconolfi@mat.uniroma1.it
Department of Mathematics, Sapienza Università di Roma, Rome, Italy

ALFONSO SORRENTINO: sorrentino@mat.uniroma2.it
Department of Mathematics, Università degli Studi di Roma "Tor Vergata", Rome, Italy

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.213

HIGH-FREQUENCY APPROXIMATION OF THE INTERIOR DIRICHLET-TO-NEUMANN MAP AND APPLICATIONS TO THE TRANSMISSION EIGENVALUES

GEORGI VODEV

We study the high-frequency behaviour of the Dirichlet-to-Neumann map for an arbitrary compact Riemannian manifold with a nonempty smooth boundary. We show that far from the real axis it can be approximated by a simpler operator. We use this fact to get new results concerning the location of the transmission eigenvalues on the complex plane. In some cases we obtain optimal transmission eigenvalue-free regions.

1. Introduction and statement of results

Let $(X, \mathcal{G})$ be a compact Riemannian manifold of dimension $d = \dim X \geq 2$ with a nonempty smooth boundary ∂X and let Δ_X denote the negative Laplace–Beltrami operator on $(X, \mathcal{G})$. Denote also by $\Delta_{\partial X}$ the negative Laplace–Beltrami operator on $(\partial X, \mathcal{G}_0)$, which is a Riemannian manifold without boundary of dimension $d-1$, where $\mathcal{G}_0$ is the Riemannian metric on ∂X induced by the metric $\mathcal{G}$. Given a function $f \in H^{m+1}(\partial X)$, let u solve

$$\begin{cases} (\Delta_X + \lambda^2 n(x))u = 0 & \text{in } X, \\ u = f & \text{on } \partial X, \end{cases} \tag{1-1}$$

where $\lambda \in \mathbb{C}$, $1 \ll |\operatorname{Im}\lambda| \ll \operatorname{Re}\lambda$ and $n \in C^\infty(\bar{X})$ is a strictly positive function. Then the Dirichlet-to-Neumann (DN) map

$$\mathcal{N}(\lambda; n) : H^{m+1}(\partial X) \to H^m(\partial X)$$

is defined by

$$\mathcal{N}(\lambda; n) f := \partial_\nu u|_{\partial X},$$

where ν is the unit inner normal to ∂X. One of our goals in the present paper is to approximate the operator $\mathcal{N}(\lambda; n)$ when $n(x) \equiv 1$ in X by a simpler one of the form $p(-\Delta_{\partial X})$ with a suitable complex-valued function $p(\sigma)$, $\sigma \geq 0$. More precisely, the function p is defined as

$$p(\sigma) = \sqrt{\sigma - \lambda^2}, \quad \operatorname{Re} p < 0.$$

Our first result is the following:

MSC2010: 35P15.
Keywords: Dirichlet-to-Neumann map, transmission eigenvalues.

Theorem 1.1. *Let $0<\epsilon<1$ be arbitrary. Then, for every $0<\delta\ll 1$ there are constants $C_\delta, C_{\epsilon,\delta}>1$ such that we have*

$$\|\mathcal{N}(\lambda;1)-p(-\Delta_{\partial X})\|_{L^2(\partial X)\to L^2(\partial X)}\le\delta|\lambda| \tag{1-2}$$

for $C_\delta\le|\operatorname{Im}\lambda|\le(\operatorname{Re}\lambda)^{1-\epsilon}$, $\operatorname{Re}\lambda\ge C_{\epsilon,\delta}$.

Note that this result has been previously proved in [Petkov and Vodev 2017b] in the case when X is a ball in $\mathbb{R}^d$ and the metric is the Euclidean one. In fact, in this case we have a better approximation of the operator $\mathcal{N}(\lambda;1)$. In the general case when the function n is arbitrary, the DN map can be approximated by h-ΨDOs, where $0<h\ll 1$ is a semiclassical parameter such that $\operatorname{Re}(h\lambda)^2=1$. To describe this more precisely let us introduce the class of symbols $S^k_\delta(\partial X)$, $0\le\delta<\frac{1}{2}$, as being the set of all functions $a(x',\xi')\in C^\infty(T^*\partial X)$ satisfying the bounds

$$\left|\partial^\alpha_{x'}\partial^\beta_{\xi'}a(x',\xi')\right|\le C_{\alpha,\beta}h^{-\delta(|\alpha|+|\beta|)}\langle\xi'\rangle^{k-|\beta|}$$

for all multi-indices α and β with constants $C_{\alpha,\beta}$ independent of h. We let $\mathrm{OP}S^k_\delta(\partial X)$ denote the set of all h-ΨDOs, $\mathrm{Op}_h(a)$, with symbol $a\in S^k_\delta(\partial X)$, defined by

$$(\mathrm{Op}_h(a)f)(x')=(2\pi h)^{-d+1}\int_{T^*\partial X}e^{-(i/h)\langle x'-y',\xi'\rangle}a(x',\xi')f(y')\,dy'\,d\xi'.$$

It is well known that for this class of symbols we have a very nice pseudodifferential calculus; e.g., see [Dimassi and Sjöstrand 1999]. It was proved in [Vodev 2015] that for $|\operatorname{Im}\lambda|\ge|\lambda|^{1/2+\epsilon}$, $0<\epsilon\ll 1$, the operator $h\mathcal{N}(\lambda;n)$ is an h-ΨDO of class $\mathrm{OP}S^1_{1/2-\epsilon}(\partial X)$ with a principal symbol

$$\rho(x',\xi')=\sqrt{r_0(x',\xi')-(h\lambda)^2n_0(x')},\quad \operatorname{Re}\rho<0,\ n_0:=n|_{\partial X},$$

$r_0\ge 0$ being the principal symbol of $-\Delta_{\partial X}$. Note that it is still possible to construct a semiclassical parametrix for the operator $h\mathcal{N}(\lambda;n)$ when $|\operatorname{Im}\lambda|\ge|\lambda|^\epsilon$, $0<\epsilon\ll 1$, if one supposes that the boundary ∂X is strictly concave; see [Vodev 2016]. This construction, however, is much more complex and one has to work with symbols belonging to much worse classes near the glancing region $\Sigma=\{(x',\xi')\in T^*\partial X: r_\sharp(x',\xi')=1\}$, where $r_\sharp=n_0^{-1}r_0$. On the other hand, it seems that no parametrix construction near Σ is possible in the important region $1\ll\text{const.}\le|\operatorname{Im}\lambda|\le|\lambda|^\epsilon$. Therefore, in the present paper we follow a different approach which consists of showing that, for arbitrary manifold X, the norm of the operator $h\mathcal{N}(\lambda;n)\mathrm{Op}_h(\chi^0_\delta)$ is $\mathcal{O}(\delta)$ for every $0<\delta\ll 1$ independent of λ, provided $|\operatorname{Im}\lambda|$ and $\operatorname{Re}\lambda$ are taken big enough (see Proposition 3.3 below). Here the function $\chi^0_\delta\in C_0^\infty(T^*\partial X)$ is supported in $\{(x',\xi')\in T^*\partial X:|r_\sharp(x',\xi')-1|\le 2\delta^2\}$ and $\chi^0_\delta=1$ in $\{(x',\xi')\in T^*\partial X:|r_\sharp(x',\xi')-1|\le\delta^2\}$ (see Section 3 for the precise definition of χ^0_δ). Theorem 1.1 is an easy consequence of the following semiclassical version.

Theorem 1.2. *Let $0<\epsilon<1$ be arbitrary. Then, for every $0<\delta\ll 1$ there are constants $C_\delta, C_{\epsilon,\delta}>1$ such that we have*

$$\left\|h\mathcal{N}(\lambda;n)-\mathrm{Op}_h(\rho(1-\chi^0_\delta)+hb)\right\|_{L^2(\partial X)\to H^1_h(\partial X)}\le C\delta \tag{1-3}$$

for $C_\delta\le|\operatorname{Im}\lambda|\le(\operatorname{Re}\lambda)^{1-\epsilon}$, $\operatorname{Re}\lambda\ge C_{\epsilon,\delta}$, where $C>0$ is a constant independent of λ and δ, and $b\in S^0_0(\partial X)$ is independent of λ and the function n.

Here $H_h^1(\partial X)$ denotes the Sobolev space equipped with the semiclassical norm (see Section 3 for the precise definition). Thus, to prove (1-3), as well as (1-2), it suffices to construct a semiclassical parametrix outside a δ^2-neighbourhood of Σ, which turns out to be much easier and can be done for an arbitrary X. In the elliptic region $\{(x',\xi') \in T^*\partial X : r_\sharp(x',\xi') \geq 1+\delta^2\}$ we use the same parametrix construction as in [Vodev 2015] with slight modifications. In the hyperbolic region $\{(x',\xi') \in T^*\partial X : r_\sharp(x',\xi') \leq 1-\delta^2\}$, however, we need to improve the parametrix construction of that paper. We do this in Section 4 for $1 \ll \text{const.} \leq |\operatorname{Im}\lambda| \leq |\lambda|^{1-\epsilon}$. Then we show that the difference between the operator $h\mathcal{N}(\lambda; n)$ microlocalized in the hyperbolic region and its parametrix is $\mathcal{O}(e^{-\beta|\operatorname{Im}\lambda|}) + \mathcal{O}_{\epsilon,M}(|\lambda|^{-M})$, where $\beta > 0$ is some constant and $M \geq 1$ is arbitrary. So, we can make it small by taking $|\operatorname{Im}\lambda|$ and $|\lambda|$ big enough.

These kinds of approximations of the DN map are important for the study of the location of the complex eigenvalues associated to boundary-value problems with dissipative boundary conditions; e.g., see [Petkov 2016]. In particular, Theorem 1.2 leads to significant improvements of the eigenvalue-free regions in that paper. In the present paper we use Theorem 1.2 to study the location of the interior transmission eigenvalues (see Section 2). We improve most of the results in [Vodev 2015], as well as those in [Petkov and Vodev 2017b; Vodev 2016], and provide simpler proofs. In some cases we get optimal transmission eigenvalue-free regions (see Theorem 2.1). Note that for the applications in the anisotropic case it suffices to have a weaker analogue of the estimate (1-3) with the space H_h^1 replaced by L^2, in which case the operator $\operatorname{Op}_h(hb)$ becomes negligible. In the isotropic case, however, it is essential to have in (1-3) the space H_h^1 and that the function b does not depend on the refraction index n.

Note finally that Theorem 1.2 can be also used to study the location of the resonances for the exterior transmission problems considered in [Cardoso et al. 2001; Galkowski 2015]. For example, it allows us to simplify the proof of the resonance-free regions in [Cardoso et al. 2001] and to extend it to more general boundary conditions.

2. Applications to the transmission eigenvalues

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded, connected domain with a C^∞ smooth boundary $\Gamma = \partial\Omega$. A complex number $\lambda \in \mathbb{C}$, $\operatorname{Re}\lambda \geq 0$, will be said to be a transmission eigenvalue if the following problem has a nontrivial solution:

$$\begin{cases} (\nabla c_1(x)\nabla + \lambda^2 n_1(x))u_1 = 0 & \text{in } \Omega, \\ (\nabla c_2(x)\nabla + \lambda^2 n_2(x))u_2 = 0 & \text{in } \Omega, \\ u_1 = u_2, \quad c_1\,\partial_\nu u_1 = c_2\,\partial_\nu u_2 & \text{on } \Gamma, \end{cases} \tag{2-1}$$

where ν denotes the Euclidean unit inner normal to Γ, $c_j, n_j \in C^\infty(\overline{\Omega})$, $j = 1, 2$, are strictly positive real-valued functions. We will consider two cases:

$$c_1(x) \equiv c_2(x) \equiv 1 \quad \text{in } \Omega, \qquad n_1(x) \neq n_2(x) \quad \text{on } \Gamma \quad \text{(isotropic case)}, \tag{2-2}$$

$$(c_1(x) - c_2(x))(c_1(x)n_1(x) - c_2(x)n_2(x)) \neq 0 \quad \text{on } \Gamma \quad \text{(anisotropic case)}. \tag{2-3}$$

In Section 6 we will prove the following:

Theorem 2.1. *Assume either the condition* (2-2) *or the condition*

$$(c_1(x) - c_2(x))\,(c_1(x)n_1(x) - c_2(x)n_2(x)) < 0 \quad \textit{on } \Gamma. \tag{2-4}$$

Then there exists a constant $C > 0$ such that there are no transmission eigenvalues in the region

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 1,\ |\operatorname{Im}\lambda| \geq C\}. \tag{2-5}$$

Remark. It is proven in [Vodev 2015] that under the condition (2-2) (as well as the condition (2-6) below) there exists a constant $\widetilde{C} > 0$ such that there are no transmission eigenvalues in the region

$$\{\lambda \in \mathbb{C} : 0 \leq \operatorname{Re}\lambda \leq 1,\ |\operatorname{Im}\lambda| \geq \widetilde{C}\}.$$

This is no longer true under the condition (2-4), in which case there exist infinitely many transmission eigenvalues very close to the imaginary axis.

Note that the eigenvalue-free region (2-5) is optimal and cannot be improved in general. Indeed, it follows from the analysis in [Leung and Colton 2012] (see Section 4) that in the isotropic case when the domain Ω is a ball and the refraction indices n_1 and n_2 are constant, there may exist infinitely many transmission eigenvalues whose imaginary parts are bounded from below by a positive constant. Note also that the above result has been previously proved in [Petkov and Vodev 2017b] in the case when the domain Ω is a ball and the coefficients are constant. In the isotropic case, the eigenvalue-free region (2-5) has been also obtained in [Sylvester 2013] when the dimension is 1. In the general case of arbitrary domains, the existence of transmission eigenvalue-free regions has been previously proved in [Hitrik et al. 2011; Lakshtanov and Vainberg 2013; Robbiano 2013] in the isotropic case, and [Vodev 2015, 2016] in both cases. For example, it has been proved in [Vodev 2015] that, under the conditions (2-2) and (2-4), there are no transmission eigenvalues in

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 1,\ |\operatorname{Im}\lambda| \geq C_\epsilon(\operatorname{Re}\lambda)^{1/2+\epsilon}\}, \quad C_\epsilon > 0,$$

for every $0 < \epsilon \ll 1$. This eigenvalue-free region has been improved in [Vodev 2016] under an additional strict concavity condition on the boundary Γ to

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 1,\ |\operatorname{Im}\lambda| \geq C_\epsilon(\operatorname{Re}\lambda)^{\epsilon}\}, \quad C_\epsilon > 0,$$

for every $0 < \epsilon \ll 1$. When the function in the left-hand side of (2-3) is strictly positive, the existence of parabolic eigenvalue-free regions has been proved in [Vodev 2015] for arbitrary domains, which however are worse than the eigenvalue-free regions we have under the conditions (2-2) and (2-4). In Section 7 we will prove:

Theorem 2.2. *Assume the conditions*

$$(c_1(x) - c_2(x))\,(c_1(x)n_1(x) - c_2(x)n_2(x)) > 0 \quad \textit{on } \Gamma \tag{2-6}$$

and

$$\frac{n_1(x)}{c_1(x)} \neq \frac{n_2(x)}{c_2(x)} \quad \textit{on } \Gamma. \tag{2-7}$$

Then there exists a constant $C > 0$ such that there are no transmission eigenvalues in the region

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 1,\ |\operatorname{Im}\lambda| \geq C\log(\operatorname{Re}\lambda + 1)\}. \tag{2-8}$$

Note that in the case when (2-6) is fulfilled but (2-7) is not, the method developed in the present paper does not work and it is not clear if improvements are possible compared with the results in [Vodev 2015]. To the best of our knowledge, no results exist in the degenerate case when the function in the left-hand side of (2-3) vanishes without being identically zero.

It has been proved in [Petkov and Vodev 2017a] that the counting function

$$N(r) = \#\{\lambda - \text{trans. eig.} :\ |\lambda| \leq r\}, \quad r > 1,$$

satisfies the asymptotics

$$N(r) = (\tau_1 + \tau_2)r^d + \mathcal{O}_\epsilon(r^{d-\kappa+\epsilon}) \quad \forall\, 0 < \epsilon \ll 1,$$

where $0 < \kappa \leq 1$ is such that there are no transmission eigenvalues in the region

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 1,\ |\operatorname{Im}\lambda| \geq C(\operatorname{Re}\lambda)^{1-\kappa}\}, \quad C > 0,$$

and

$$\tau_j = \frac{\omega_d}{(2\pi)^d}\int_\Omega \left(\frac{n_j(x)}{c_j(x)}\right)^{d/2} dx,$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$. Using this we obtain from the above theorems the following:

Corollary 2.3. *Under the conditions of Theorems 2.1 and 2.2, the counting function of the transmission eigenvalues satisfies the asymptotics*

$$N(r) = (\tau_1 + \tau_2)r^d + \mathcal{O}_\epsilon(r^{d-1+\epsilon}) \quad \forall\, 0 < \epsilon \ll 1. \tag{2-9}$$

This result has been previously proved in [Vodev 2016] under an additional strict concavity condition on the boundary Γ. In the present paper we remove this additional condition to conclude that in fact the asymptotics (2-9) holds true for an arbitrary domain. We also expect that (2-9) holds with $\epsilon = 0$, but this remains an interesting open problem. In the isotropic case asymptotics for the counting function $N(r)$ with remainder $o(r^d)$ have been previously obtained in [Faierman 2014; Pham and Stefanov 2014; Robbiano 2016].

3. A priori estimates in the glancing region

Let $\lambda \in \mathbb{C}$, $\operatorname{Re}\lambda > 1$, $1 < |\operatorname{Im}\lambda| \leq \theta_0 \operatorname{Re}\lambda$, where $0 < \theta_0 < 1$ is a fixed constant, and set $h = \mu^{-1}$, where

$$\mu = \operatorname{Re}\lambda\sqrt{1 - \left(\frac{\operatorname{Im}\lambda}{\operatorname{Re}\lambda}\right)^2} \sim \operatorname{Re}\lambda \sim |\lambda|.$$

Clearly, we have $\operatorname{Re}(h\lambda)^2 = 1$ and

$$\lambda^2 = \mu^2(1 + izh), \quad z = 2\mu^{-1}\operatorname{Im}\lambda\operatorname{Re}\lambda \sim 2\operatorname{Im}\lambda.$$

Given an integer $m \geq 0$, denote by $H_h^m(X)$ the Sobolev space equipped with the semiclassical norm

$$\|v\|_{H_h^m(X)} = \sum_{|\alpha|\leq m} h^{|\alpha|} \|\partial_x^\alpha v\|_{L^2(X)}.$$

We define similarly the Sobolev space $H_h^m(\partial X)$. It is well known that

$$\|v\|_{H_h^m(\partial X)} \sim \|\mathrm{Op}_h(\langle \xi' \rangle^m) v\|_{L^2(\partial X)} \sim \|v\|_{L^2(\partial X)} + \|\mathrm{Op}_h((1-\eta)|\xi'|^m) v\|_{L^2(\partial X)}$$

for any function $\eta \in C_0^\infty(T^*\partial X)$ independent of h. Hereafter, $\langle \xi' \rangle = (1+|\xi'|^2)^{1/2}$.

Given functions $V \in L^2(X)$ and $f \in L^2(\partial X)$, we let the function u solve

$$\begin{cases} (\Delta_X + \lambda^2 n(x))u = \lambda V & \text{in } X, \\ u = f & \text{on } \partial X, \end{cases} \tag{3-1}$$

and set $g = h\, \partial_\nu u|_{\partial X}$. We will first prove:

Lemma 3.1. *There is a constant $C > 0$ such that the following estimate holds*:

$$\|u\|_{H_h^1(X)} \leq C|\mathrm{Im}\,\lambda|^{-1} \|V\|_{L^2(X)} + C|\mathrm{Im}\,\lambda|^{-1/2} \|f\|_{L^2(\partial X)}^{1/2} \|g\|_{L^2(\partial X)}^{1/2}. \tag{3-2}$$

Proof. By Green's formula we have

$$\mathrm{Im}(\lambda^2)\|n^{1/2}u\|_{L^2(X)}^2 = \mathrm{Im}\langle \lambda V, u\rangle_{L^2(X)} + \mathrm{Im}\langle \partial_\nu u|_{\partial X}, f\rangle_{L^2(\partial X)}$$

which implies

$$|\mathrm{Im}\,\lambda|\, \|u\|_{L^2(X)}^2 \lesssim \|V\|_{L^2(X)} \|u\|_{L^2(X)} + \|f\|_{L^2(\partial X)} \|g\|_{L^2(\partial X)}. \tag{3-3}$$

On the other hand, we have

$$\|\nabla_X u\|_{L^2(X)}^2 - \mathrm{Re}(\lambda^2)\|n^{1/2}u\|_{L^2(X)}^2 = -\mathrm{Re}\langle \lambda V, u\rangle_{L^2(X)} - \mathrm{Re}\langle \partial_\nu u|_{\partial X}, f\rangle_{L^2(\partial X)},$$

which yields

$$\|h\nabla_X u\|_{L^2(X)}^2 \lesssim \|u\|_{L^2(X)}^2 + \mathcal{O}(h^2)\|V\|_{L^2(X)}^2 + \mathcal{O}(h)\|f\|_{L^2(\partial X)} \|g\|_{L^2(\partial X)}. \tag{3-4}$$

Since $h \lesssim |\mathrm{Im}\,\lambda|^{-1}$, the estimate (3-2) follows from (3-3) and (3-4). □

We now equip X with the Riemannian metric $n\mathcal{G}$. We will write the operator $n^{-1}\Delta_X$ in the normal coordinates (x_1, x') with respect to the metric $n\mathcal{G}$ near the boundary ∂X, where $0 < x_1 \ll 1$ denotes the distance to the boundary and x' are coordinates on ∂X. Set $\Gamma(x_1) = \{x \in X : \mathrm{dist}(x, \partial X) = x_1\}$, $\Gamma(0) = \partial X$. Then $\Gamma(x_1)$ is a Riemannian manifold without boundary of dimension $d-1$ with a Riemannian metric induced by the metric $n\mathcal{G}$, which depends smoothly in x_1. It is well known that the operator $n^{-1}\Delta_X$ can be written as

$$n^{-1}\Delta_X = \partial_{x_1}^2 + Q(x_1) + R,$$

where $Q(x_1) = \Delta_{\Gamma(x_1)}$ is the negative Laplace–Beltrami operator on $\Gamma(x_1)$ and R is a first-order differential operator. Clearly, $Q(x_1)$ is a second-order differential operator with smooth coefficients and $Q(0) = \Delta_{\partial X}^{(n)}$ is the negative Laplace–Beltrami operator on ∂X equipped with the Riemannian metric induced by the metric $n\mathcal{G}$.

Let $\chi \in C_0^\infty(\mathbb{R})$, $0 \le \chi(t) \le 1$, $\chi(t) = 1$ for $|t| \le 1$ and $\chi(t) = 0$ for $|t| \ge 2$. Given a parameter $0 < \delta_1 \ll 1$ independent of λ and an integer $k \ge 0$, set $\phi_k(x_1) = \chi(2^{-k}x_1/\delta_1)$. Given integers $0 \le s_1 \le s_2$, we define the norm $\|u\|_{s_1,s_2,k}$ by

$$\|u\|^2_{s_1,s_2,k} = \|u\|^2_{H_h^{s_1}(X)} + \sum_{\ell_1=0}^{s_1} \sum_{\ell_2=0}^{s_2-\ell_1} \int_0^\infty \|(h\,\partial_{x_1})^{\ell_1}(\phi_k u)(x_1,\cdot)\|^2_{H_h^{\ell_2}(\partial X)}\,dx_1.$$

Clearly, we have

$$\|u\|_{H_h^{s_1}(X)} \le \|u\|_{s_1,s_2,k} \lesssim \|u\|_{H_h^{s_2}(X)}.$$

Throughout this paper $\eta \in C_0^\infty(T^*\partial X)$, $0 \le \eta \le 1$, $\eta = 1$ in $|\xi'| \le A$, $\eta = 0$ in $|\xi'| \ge A+1$, will be a function independent of λ, where $A > 1$ is a parameter we may take as large as we want. We will now prove:

Lemma 3.2. *Let u solve* (3-1) *with $V \in H^{s-1}(X)$ and $f \in H^{2s}(\partial X)$ for some integer $s \ge 1$. Then the following estimate holds*:

$$\|u\|_{1,s+1,k} \lesssim \|u\|_{H_h^1(X)} + \|V\|_{0,s-1,k+s-1} + \|\mathrm{Op}_h(1-\eta)f\|^{1/2}_{H_h^{2s}(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}. \tag{3-5}$$

Proof. Note that

$$\|u\|_{1,s+1,k} \lesssim \|u\|_{H_h^1(X)} + \|u_{s,k}\|_{H_h^1(X)},$$

where the function $u_{s,k} = \mathrm{Op}_h((1-\eta)|\xi'|^s)(\phi_k u)$ satisfies the equation

$$(h^2\partial_{x_1}^2 + h^2 Q(x_1) + 1 + ihz)u_{s,k} = U_{s,k}$$

with

$$\begin{aligned} U_{s,k} = \big[h^2 Q(x_1), \mathrm{Op}_h((1-\eta)|\xi'|^s)\big](\phi_k u) + \mathrm{Op}_h((1-\eta)|\xi'|^s)[h^2\partial_{x_1}^2, \phi_k]\phi_{k+1}u \\ - h^2\mathrm{Op}_h((1-\eta)|\xi'|^s)\phi_k R\phi_{k+1}u + h^2\lambda\mathrm{Op}_h((1-\eta)|\xi'|^s)(\phi_k V). \end{aligned}$$

We also have

$$\begin{aligned} f_s &:= u_{s,k}|_{x_1=0} = \mathrm{Op}_h((1-\eta)|\xi'|^s)f, \\ g_s &:= h\,\partial_{x_1}u_{s,k}|_{x_1=0} = \mathrm{Op}_h((1-\eta)|\xi'|^s)g_\flat, \end{aligned}$$

where $g_\flat := h\,\partial_{x_1}u|_{x_1=0}$. Integrating by parts the above equation and taking the real part, we get

$$\begin{aligned} &\|h\,\partial_{x_1}u_{s,k}\|^2_{L^2(X)} - \langle (h^2Q(x_1)+1)u_{s,k}, u_{s,k}\rangle_{L^2(X)} \\ &\qquad \le |\langle U_{s,k}, u_{s,k}\rangle_{L^2(X)}| + h|\langle f_s, g_s\rangle_{L^2(\partial X)}| \\ &\qquad \lesssim \|u_{s,k}\|_{H_h^1(X)}(\|V\|_{0,s-1,k} + \|u\|_{1,s,k+1}) \\ &\qquad\qquad + \big\|\mathrm{Op}_h((1-\eta)|\xi'|^s)^*\mathrm{Op}_h((1-\eta)|\xi'|^s)f\big\|_{L^2(\partial X)}\|g_\flat\|_{L^2(\partial X)}. \end{aligned} \tag{3-6}$$

The principal symbol r of the operator $-Q(x_1)$ satisfies $r(x,\xi') \ge C'|\xi'|^2$, $C' > 0$, on $\mathrm{supp}\,\phi_k$, provided δ_1 is taken small enough. Therefore, we can arrange by taking the parameter A big enough that $r - 1 \ge C\langle\xi'\rangle$ on $\mathrm{supp}(1-\eta)\phi_k$, where $C > 0$ is some constant. Hence, by Gårding's inequality we have

$$-\langle (h^2Q(x_1)+1)u_{s,k}, u_{s,k}\rangle_{L^2(X)} \ge C\|\mathrm{Op}_h(\langle\xi'\rangle)u_{s,k}\|^2_{L^2(X)} \tag{3-7}$$

with possibly a new constant $C > 0$. Since the norms of g and $g_\flat$ are equivalent, by (3-6) and (3-7) we get

$$\|u_{s,k}\|_{H_h^1(X)} \lesssim \|V\|_{0,s-1,k} + \|u\|_{H_h^1(X)} + \|u_{s-1,k+1}\|_{H_h^1(X)} + \|\mathrm{Op}_h(1-\eta)f\|^{1/2}_{H_h^{2s}(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}. \tag{3-8}$$

We may now apply the same argument to $u_{s-1,k+1}$. Thus, repeating this argument a finite number of times we can eliminate the term involving $u_{s-1,k+1}$ in the right-hand side of (3-8) and obtain the estimate (3-5). □

Let the functions $\chi_j \in C^\infty(\mathbb{R})$, $0 \le \chi_j(t) \le 1$, $j = 1, 2, 3$, be such that $\chi_1 + \chi_2 + \chi_3 \equiv 1$, $\chi_2 = \chi$, $\chi_1(t) = 1$ for $t \le -2$, $\chi_1(t) = 0$ for $t \ge -1$, $\chi_3(t) = 0$ for $t \le 1$, $\chi_3(t) = 1$ for $t \ge 2$. Given a parameter $0 < \delta \ll 1$ independent of λ, set

$$\begin{aligned} \chi_\delta^-(x',\xi') &= \chi_1\big((r_\sharp(x',\xi')-1)/\delta^2\big),\\ \chi_\delta^0(x',\xi') &= \chi_2\big((r_\sharp(x',\xi')-1)/\delta^2\big),\\ \chi_\delta^+(x',\xi') &= \chi_3\big((r_\sharp(x',\xi')-1)/\delta^2\big),\end{aligned}$$

where $r_\sharp = n_0^{-1} r_0$ is the principal symbol of the operator $-\Delta_{\partial X}^{(n)}$. Since $(r_\sharp - 1)^k \chi_\delta^0 = \mathcal{O}(\delta^{2k})$, we have

$$(h^2\Delta_{\partial X}^{(n)} + 1)^k \mathrm{Op}_h(\chi_\delta^0) = \mathcal{O}(\delta^{2k}) : L^2(\partial X) \to L^2(\partial X) \tag{3-9}$$

for every integer $k \ge 0$. Clearly, we also have

$$\mathrm{Op}_h(\chi_\delta^0) = \mathcal{O}(1) : L^2(\partial X) \to H_h^m(\partial X) \quad \forall\, m \ge 0,$$

uniformly in δ. Using (3-9) we will prove:

Proposition 3.3. *Let u solve* (3-1) *with $f \equiv 0$ and $V \in H^s(X)$ for some integer $s \ge 0$. Then the function $g = h\,\partial_\nu u|_{\partial X}$ satisfies the estimate*

$$\|g\|_{H_h^s(\partial X)} \le C'|\mathrm{Im}\,\lambda|^{-1/2}\|V\|_{0,s,s} \tag{3-10}$$

with a constant $C' > 0$ independent of λ.

Let u solve (3-1) *with f replaced by $\mathrm{Op}_h(\chi_\delta^0) f$ and $V \in H^{s+2}(X)$ for some integer $s \ge 0$. Then the function $g = h\,\partial_\nu u|_{\partial X}$ satisfies the estimate*

$$\|g\|_{H_h^s(\partial X)} \le C(\delta + |\mathrm{Im}\,\lambda|^{-1/4})\|f\|_{L^2(\partial X)} + C(\delta^{1/2} + |\mathrm{Im}\,\lambda|^{-1/8})\|V\|_{0,s+2,s+2} \tag{3-11}$$

for $1 < |\mathrm{Im}\,\lambda| \le \delta^2\,\mathrm{Re}\,\lambda$, $\mathrm{Re}\,\lambda \ge C_\delta \gg 1$, with a constant $C > 0$ independent of λ and δ.

Proof. Set $w = \phi_0(x_1)u$. We will first show that the estimates (3-10) and (3-11) with $s \ge 1$ follow from (3-10) and (3-11) with $s = 0$, respectively. This follows from the estimate

$$\|g\|_{H_h^s(\partial X)} \lesssim \|g\|_{L^2(\partial X)} + \big\|h\,\partial_{x_1} v_s|_{x_1=0}\big\|_{L^2(\partial X)}, \tag{3-12}$$

where the function $v_s = \mathrm{Op}_h((1-\eta)|\xi'|^s)w$ satisfies (3-1) with V replaced by

$$V_s = n\mathrm{Op}_h((1-\eta)|\xi'|^s)\phi_0 n^{-1}V + \lambda^{-1}n\big[n^{-1}\Delta_X, \mathrm{Op}_h((1-\eta)|\xi'|^s)\phi_0\big]u.$$

We can write the commutator as

$$[\partial_{x_1}^2 + R, \phi_0(x_1)]\mathrm{Op}_h((1-\eta)|\xi'|^s)\phi_1(x_1) + \phi_0\big[Q(x_1) + R, \mathrm{Op}_h((1-\eta)|\xi'|^s)\big]\phi_1(x_1).$$

Therefore, if $f \equiv 0$, in view of Lemmas 3.1 and 3.2, the function V_s satisfies the bound

$$\|V_s\|_{0,0,0} \lesssim \|V\|_{0,s,0} + \|u\|_{1,s+1,1} \lesssim \|u\|_{H_h^1(X)} + \|V\|_{0,s,s} \lesssim \|V\|_{0,s,s}. \tag{3-13}$$

Clearly, the assertion concerning (3-10) follows from (3-12) and (3-13). The estimate (3-11) can be treated similarly. Indeed, in view of Lemma 3.2, the function V_s satisfies the bound

$$\begin{aligned}\|V_s\|_{0,2,2} &\lesssim \|V\|_{0,s+2,0} + \|u\|_{1,s+3,1} \\ &\lesssim \|u\|_{H_h^1(X)} + \|V\|_{0,s+2,s+2} + \|\mathrm{Op}_h(1-\eta)\mathrm{Op}_h(\chi_\delta^0) f\|^{1/2}_{H_h^{2s+4}(\partial X)} \|g\|^{1/2}_{L^2(\partial X)}.\end{aligned} \tag{3-14}$$

Taking the parameter A big enough we can arrange that $\operatorname{supp} \chi_\delta^0 \cap \operatorname{supp}(1-\eta) = \varnothing$. Hence

$$\mathrm{Op}_h(1-\eta)\mathrm{Op}_h(\chi_\delta^0) = \mathcal{O}(h^\infty) : L^2(\partial X) \to H_h^m(\partial X) \quad \forall m \geq 0. \tag{3-15}$$

By (3-14) and (3-15) together with Lemma 3.1 we conclude

$$\begin{aligned}\|V_s\|_{0,2,2} &\lesssim \|u\|_{H_h^1(X)} + \|V\|_{0,s+2,s+2} + \mathcal{O}(h^\infty)\|f\|^{1/2}_{L^2(\partial X)}\|g\|^{1/2}_{L^2(\partial X)} \\ &\lesssim \|V\|_{0,s+2,s+2} + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2} + h^\infty)\|f\|^{1/2}_{L^2(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}.\end{aligned}$$

We now apply (3-11) with $s = 0$ to the function v_s and note that

$$v_s|_{x_1=0} = \mathrm{Op}_h((1-\eta)|\xi'|^s)\mathrm{Op}_h(\chi_\delta^0) f = \mathcal{O}(h^\infty) f.$$

Hence

$$\begin{aligned}\big\|h\,\partial_{x_1} v_s|_{x_1=0}\big\|_{L^2(\partial X)} &\leq \mathcal{O}(h^\infty)\|f\|_{L^2(\partial X)} + \mathcal{O}(\delta^{1/2} + |\mathrm{Im}\,\lambda|^{-1/8})\|V_s\|_{0,2,2} \\ &\leq \mathcal{O}(\delta^{1/2} + |\mathrm{Im}\,\lambda|^{-1/8})\|V\|_{0,s+2,s+2} + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2} + h^\infty)\|f\|^{1/2}_{L^2(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}.\end{aligned} \tag{3-16}$$

Therefore, the assertion concerning (3-11) follows from (3-12) and (3-16).

We now turn to the proofs of (3-10) and (3-11) with $s = 0$. In view of Lemma 3.1, the function

$$U := h(n^{-1}\Delta_X + \lambda^2)w = h[n^{-1}\Delta_X, \phi_0(x_1)]u + h\lambda n^{-1}\phi_0 V$$

satisfies the bound

$$\|U\|_{L^2(X)} \lesssim \|u\|_{H_h^1(X)} + \|V\|_{L^2(X)} \lesssim \|V\|_{L^2(X)} + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2})\|f\|^{1/2}_{L^2(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}. \tag{3-17}$$

Observe now that the derivative of the function

$$E(x_1) = \|h\,\partial_{x_1} w\|^2 + \langle (h^2 Q(x_1) + 1)w, w\rangle,$$

where $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ are the norm and the scalar product in $L^2(\partial X)$, satisfies

$$\begin{aligned}E'(x_1) &= 2\,\mathrm{Re}\big\langle (h^2\partial_{x_1}^2 + h^2 Q(x_1) + 1)w, \partial_{x_1} w\big\rangle + \langle h^2 Q'(x_1)w, w\rangle \\ &= 2\,\mathrm{Re}\big\langle (U - izw - hRw), h\,\partial_{x_1} w\big\rangle + \langle h^2 Q'(x_1)w, w\rangle.\end{aligned}$$

If we put $g_\flat := h\,\partial_{x_1} u|_{x_1=0}$, we have

$$\begin{aligned}\|g_\flat\|^2 + \big\langle (h^2\Delta_{\partial X}^{(n)}+1)\mathrm{Op}_h(\chi_\delta^0)f, \mathrm{Op}_h(\chi_\delta^0)f\big\rangle \\ &= E(0) = -\int_0^\infty E'(x_1)\,dx_1 \\ &\lesssim \big(\|U\|_{L^2(X)} + |z|\,\|w\|_{L^2(X)} + \|hRw\|_{L^2(X)}\big)\|h\,\partial_{x_1}w\|_{L^2(X)} + \|w\|^2_{H_h^1(X)} \\ &\le \mathcal{O}(|z|)\|h\,\partial_{x_1}w\|_{L^2(X)}\|w\|_{L^2(X)} + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1})F^2, \qquad (3\text{-}18)\end{aligned}$$

where we have used Lemma 3.1 together with (3-17) and we have put

$$F = \|f\|^{1/2}\|g\|^{1/2} + \|V\|_{L^2(X)}.$$

Clearly, (3-10) with $s = 0$ follows from (3-18) applied with $f \equiv 0$ and Lemma 3.1. To prove (3-11) with $s = 0$, observe that (3-9) and (3-18) lead to

$$\|g\| \le \mathcal{O}(\delta)\|f\| + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2})F + \mathcal{O}(|\mathrm{Im}\,\lambda|^{1/2})\|h\,\partial_{x_1}w\|^{1/2}_{L^2(X)}\|w\|^{1/2}_{L^2(X)}. \qquad (3\text{-}19)$$

We now need a better bound on the norm $\|h\,\partial_{x_1}w\|_{L^2(X)}$ in the right-hand side of (3-19) than what the estimate (3-2) gives. To this end, observe that integrating by parts yields

$$\begin{aligned}\|h\,\partial_{x_1}w\|^2_{L^2(X)} - \langle (h^2Q(x_1)+1)w, w\rangle_{L^2(X)} &= -h\,\mathrm{Re}\langle (U-hRw), w\rangle_{L^2(X)} - h\,\mathrm{Re}\langle f, g_\flat\rangle \\ &\le \mathcal{O}(h)\|w\|^2_{H_h^1(X)} + \mathcal{O}(h)\|U\|^2_{L^2(X)} + \mathcal{O}(h)\|f\|\,\|g\| \\ &\le \mathcal{O}(h)F^2. \qquad (3\text{-}20)\end{aligned}$$

By (3-19) and (3-20), together with Lemma 3.1, we get

$$\begin{aligned}\|g\| &\le \mathcal{O}(\delta)\|f\| + \mathcal{O}(|\mathrm{Im}\,\lambda|^{1/2})\|w_1\|^{1/4}_{L^2(X)}\|w\|^{3/4}_{L^2(X)} + \mathcal{O}(h^{1/4}|\mathrm{Im}\,\lambda|^{1/2})F^{1/2}\|w\|^{1/2}_{L^2(X)} + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2})F \\ &\le \mathcal{O}(\delta)\|f\| + \mathcal{O}(|\mathrm{Im}\,\lambda|^{1/8})\|w_1\|^{1/4}_{L^2(X)}F^{3/4} + \mathcal{O}\big(|\mathrm{Im}\,\lambda|^{-1/2} + h^{1/4}|\mathrm{Im}\,\lambda|^{1/4}\big)F, \qquad (3\text{-}21)\end{aligned}$$

where we have put $w_1 := (h^2Q(x_1)+1)w$. We need now the following:

Lemma 3.4. *The function w_1 satisfies the estimate*

$$|\mathrm{Im}\,\lambda|^{1/2}\|w_1\|_{L^2(X)} \le \mathcal{O}(\delta^2 + |\mathrm{Im}\,\lambda|^{-1} + h^\infty)\|f\|^{1/2}\|g\|^{1/2} + \mathcal{O}(h^{1/2})\|f\| + \mathcal{O}(|\mathrm{Im}\,\lambda|^{-1/2})\|V\|_{0,2,2}. \quad (3\text{-}22)$$

Let us show that this lemma implies the estimate (3-11) with $s = 0$. Set

$$\widetilde{F} = \|f\|^{1/2}\|g\|^{1/2} + \|V\|_{0,2,2} \ge F.$$

By (3-21) and (3-22),

$$\begin{aligned}\|g\| &\le \mathcal{O}(\delta)\|f\| + \mathcal{O}(\delta^{1/2} + |\mathrm{Im}\,\lambda|^{-1/8} + h^\infty)\widetilde{F} + \mathcal{O}(h^{1/8})(\|f\| + F) + \mathcal{O}\big(|\mathrm{Im}\,\lambda|^{-1/2} + h^{1/4}|\mathrm{Im}\,\lambda|^{1/4}\big)F \\ &\le \mathcal{O}(\delta + h^{1/8})\|f\| + \mathcal{O}\big(\delta^{1/2} + |\mathrm{Im}\,\lambda|^{-1/8} + h^{1/8} + h^{1/4}|\mathrm{Im}\,\lambda|^{1/4}\big)\widetilde{F}. \qquad (3\text{-}23)\end{aligned}$$

Since by assumption $h^{1/4}|\mathrm{Im}\,\lambda|^{1/4} = \mathcal{O}(\delta^{1/2})$, one can easily see that (3-11) with $s = 0$ follows from (3-23). $\square$

Proof of Lemma 3.4. Observe that the function w_1 satisfies the equation

$$(h^2\partial_{x_1}^2 + h^2 Q(x_1) + 1 + ihz)w_1 = hU_1,$$

where

$$U_1 := (h^2 Q(x_1)+1)(U - hRw) + 2h^3 Q'(x_1)\partial_{x_1} w + h^3 Q''(x_1)w.$$

We also have

$$f_1 := w_1|_{x_1=0} = (h^2 Q(0)+1)\mathrm{Op}_h(\chi_\delta^0)f,$$
$$g_1 := h\,\partial_{x_1} w_1|_{x_1=0} = (h^2 Q(0)+1)g_\flat + h^3 Q'(0)\mathrm{Op}_h(\chi_\delta^0)f.$$

Integrating by parts the above equation and taking the imaginary part, we get

$$\begin{aligned}
|z|\,\|w_1\|^2_{L^2(X)} &\le |\langle U_1, w_1\rangle_{L^2(X)}| + |\langle f_1, g_1\rangle| \\
&\le \|U_1\|_{L^2(X)}\|w_1\|_{L^2(X)} + \mathcal{O}(1)\|(h^2 Q(0)+1)^2\mathrm{Op}_h(\chi_\delta^0)f\|\,\|g\| \\
&\qquad\qquad + \mathcal{O}(h)\|\mathrm{Op}_h(\chi_\delta^0)f\|_{H_h^2(\partial X)}\|(h^2 Q(0)+1)\mathrm{Op}_h(\chi_\delta^0)f\| \\
&\le \|U_1\|_{L^2(X)}\|w_1\|_{L^2(X)} + \mathcal{O}(\delta^4)\|f\|\,\|g\| + \mathcal{O}(h)\|f\|^2,
\end{aligned}$$

where we have used (3-9). Hence

$$|z|\|w_1\|^2_{L^2(X)} \le \mathcal{O}(|z|^{-1})\|U_1\|^2_{L^2(X)} + \mathcal{O}(\delta^4)\|f\|\|g\| + \mathcal{O}(h)\|f\|^2. \tag{3-24}$$

Recall that the function U is of the form $(2h\,\partial_{x_1} + a(x))\phi_1(x_1)u + h\lambda n^{-1}\phi_0 V$, where a is some smooth function. Hence the function U_1 satisfies the estimate

$$\|U_1\|_{L^2(X)} \lesssim \|u\|_{1,3,1} + \|V\|_{0,2,0} \lesssim \|u\|_{H_h^1(X)} + \|V\|_{0,2,2} + \mathcal{O}(h^\infty)\|f\|^{1/2}_{L^2(\partial X)}\|g\|^{1/2}_{L^2(\partial X)}, \tag{3-25}$$

where we have used Lemma 3.2 together with (3-15). By (3-24) and (3-25),

$$|z|\,\|w_1\|^2_{L^2(X)} \le \mathcal{O}(|z|^{-1})\|u\|^2_{H_h^1(X)} + \mathcal{O}(|z|^{-1})\|V\|^2_{0,2,2} + \mathcal{O}(\delta^4 + h^\infty)\|f\|\,\|g\| + \mathcal{O}(h)\|f\|^2. \tag{3-26}$$

Clearly, (3-22) follows from (3-26) and Lemma 3.1. □

4. Parametrix construction in the hyperbolic region

Let λ be as in Theorems 1.1 and 1.2, and let h, z, δ, r_0, n_0, $r_\sharp$, χ and χ_δ^- be as in the previous sections. Set $\theta = \mathrm{Im}(h\lambda)^2 = hz = \mathcal{O}(h^\epsilon)$, $|\theta| \gg h$, and

$$\rho(x',\xi') = \sqrt{r_0(x',\xi') - (1+i\theta)n_0(x')}, \qquad \mathrm{Re}\,\rho < 0.$$

It is easy to see that $\rho\chi_\delta^- \in S_0^0(\partial X)$. In this section we will prove:

Proposition 4.1. *There are constants $C, C_1 > 0$ depending on δ but independent of λ such that*

$$\left\| h\mathcal{N}(\lambda; n)\mathrm{Op}_h(\chi_\delta^-) - \mathrm{Op}_h(\rho\chi_\delta^-)\right\|_{L^2(\partial X)\to H_h^1(\partial X)} \le C_1(h + e^{-C|\mathrm{Im}\,\lambda|}). \tag{4-1}$$

Proof. To prove (4-1) we will build a parametrix near the boundary of the solution to (1-1) with f replaced by $\mathrm{Op}_h(\chi_\delta^-)f$. Let $x = (x_1, x')$, $x_1 > 0$, be the normal coordinates with respect to the metric $\mathcal{G}$, which of course are different from those introduced in the previous section. In these coordinates the operator Δ_X is given by

$$\Delta_X = \partial_{x_1}^2 + \widetilde{Q} + \widetilde{R},$$

where $\widetilde{Q} \le 0$ is a second-order differential operator with respect to the variable x' and $\widetilde{R}$ is a first-order differential operator with respect to the variable x, both with coefficients depending smoothly on x. Let $(x^0, \xi^0) \in \operatorname{supp} \chi_\delta^-$ and let $\mathcal{U} \subset T^*\partial X$ be a small open neighbourhood of (x^0, ξ^0) contained in $\{r_\sharp \le 1 - \delta^2/2\}$. Take a function $\psi \in C_0^\infty(\mathcal{U})$. We will construct a parametrix $\tilde{u}_\psi^-$ of the solution of (1-1) with $\tilde{u}_\psi^-|_{x_1=0} = \mathrm{Op}_h(\psi) f$ in the form $\tilde{u}_\psi^- = \phi(x_1)\mathcal{K}^- f$, where $\phi(x_1) = \chi(x_1/\delta_1)$, $0 < \delta_1 \ll 1$, is a parameter independent of λ to be fixed later on depending on δ, and

$$(\mathcal{K}^- f)(x) = (2\pi h)^{-d+1} \iint e^{(i/h)(\langle y', \xi' \rangle + \varphi(x, \xi', \theta))} a(x, \xi', \lambda) f(y') \, d\xi' \, dy'.$$

The phase φ is complex-valued such that $\varphi|_{x_1=0} = -\langle x', \xi' \rangle$ and satisfies the eikonal equation mod $\mathcal{O}(\theta^M)$:

$$(\partial_{x_1}\varphi)^2 + \langle B(x)\nabla_{x'}\varphi, \nabla_{x'}\varphi \rangle - (1 + i\theta)n(x) = \theta^M \mathcal{R}_M, \tag{4-2}$$

where $M \gg 1$ is an arbitrary integer, the function $\mathcal{R}_M$ is bounded uniformly in θ, and B is a matrix-valued function such that $r(x, \xi') = \langle B(x)\xi', \xi' \rangle$, $r(x, \xi') \ge 0$, is the principal symbol of the operator $-\widetilde{Q}$. We clearly have $r_0(x', \xi') = r(0, x', \xi')$. Let us see that for $(x', \xi') \in \mathcal{U}$, $0 \le x_1 \le 3\delta_1$, (4-2) has a smooth solution satisfying

$$\partial_{x_1}\varphi|_{x_1=0} = -i\rho + \mathcal{O}(\theta^{M/2}) \tag{4-3}$$

provided δ_1 and $\mathcal{U}$ are small enough. We will be looking for φ in the form

$$\varphi = \sum_{j=0}^{M-1} (i\theta)^j \varphi_j(x, \xi'),$$

where φ_j are real-valued functions depending only on the sign of θ and satisfying the equations

$$(\partial_{x_1}\varphi_0)^2 + \langle B(x)\nabla_{x'}\varphi_0, \nabla_{x'}\varphi_0 \rangle = n(x), \tag{4-4}$$

$$\sum_{j=0}^{k} \partial_{x_1}\varphi_j \, \partial_{x_1}\varphi_{k-j} + \sum_{j=0}^{k} \langle B(x)\nabla_{x'}\varphi_j, \nabla_{x'}\varphi_{k-j} \rangle = \epsilon_k n(x), \quad 1 \le k \le M-1, \tag{4-5}$$

$\varphi_0|_{x_1=0} = -\langle x', \xi' \rangle$, $\varphi_j|_{x_1=0} = 0$ for $j \ge 1$, where $\epsilon_1 = 1$, $\epsilon_k = 0$ for $k \ge 2$. It is easy to check that with this choice the function φ satisfies (4-2) with $\mathcal{R}_M$ being polynomial in θ.

Clearly, if φ_0 is a solution to (4-4), then we have $(\partial_{x_1}\varphi_0|_{x_1=0})^2 = n_0(x') - r_0(x', \xi') \ge C'$ with some constant $C' > 0$ depending on δ. It is well known that (4-4) has a local (that is, for δ_1 and $\mathcal{U}$ small enough) real-valued solution $\varphi_0^\pm$ such that $\partial_{x_1}\varphi_0^\pm|_{x_1=0} = \pm\sqrt{n_0 - r_0}$. We now define the function φ_0 by $\varphi_0 = \varphi_0^+$ if $\theta > 0$ and $\varphi_0 = \varphi_0^-$ if $\theta < 0$. Hence $|\partial_{x_1}\varphi_0(x, \xi')| \ge \text{const.} > 0$ for x_1 small enough. Therefore, the equations (4-5) can be solved locally. Taking $x_1 = 0$ in (4-5) with $k = 1$, we find

$$\theta\partial_{x_1}\varphi_1|_{x_1=0} = \theta n_0 (2\partial_{x_1}\varphi_0|_{x_1=0})^{-1} = \tfrac{1}{2}|\theta| n_0 (n_0 - r_0)^{-1/2} \ge \tfrac{1}{2}C|\theta| \tag{4-6}$$

on $\mathcal{U}$, where $C = \min \sqrt{n_0(x')}$. Hence

$$\operatorname{Im} \partial_{x_1}\varphi|_{x_1=0} = \theta\partial_{x_1}\varphi_1|_{x_1=0} + \mathcal{O}(\theta^2) \geq \tfrac{1}{3}C|\theta| \tag{4-7}$$

if $|\theta|$ is taken small enough. On the other hand, taking $x_1 = 0$ in (4-2) we find

$$(\partial_{x_1}\varphi|_{x_1=0})^2 = (i\rho)^2 + \mathcal{O}(\theta^M) = (i\rho)^2(1 + \mathcal{O}(\theta^M)), \tag{4-8}$$

where we have used that $|\rho| \geq \text{const.} > 0$ on $\mathcal{U}$. Since $\operatorname{Re}\rho < 0$, we get (4-3) from (4-7) and (4-8). By (4-6) we also get

$$\theta\varphi_1(x_1, x', \xi') = \theta x_1\partial_{x_1}\varphi_1(0, x', \xi') + \mathcal{O}(\theta x_1^2) \geq \tfrac{1}{2}Cx_1|\theta| - \mathcal{O}(|\theta|x_1^2) \geq \tfrac{1}{3}Cx_1|\theta|$$

provided x_1 is taken small enough. This implies

$$\operatorname{Im} \varphi(x, \xi', \theta) = \theta\varphi_1(x_1, x', \xi') + \mathcal{O}(\theta^2 x_1) \geq \tfrac{1}{4}Cx_1|\theta|. \tag{4-9}$$

The amplitude a is of the form

$$a = \sum_{k=0}^{m} h^k a_k(x, \xi', \theta),$$

where $m \gg 1$ is an arbitrary integer and the functions a_k satisfy the transport equations mod $\mathcal{O}(\theta^M)$:

$$2i\partial_{x_1}\varphi\partial_{x_1}a_k + 2i\langle B(x)\nabla_{x'}\varphi, \nabla_{x'}a_k\rangle + i(\Delta_X\varphi)a_k + \Delta_X a_{k-1} = \theta^M \mathcal{Q}_M^{(k)}, \quad 0 \leq k \leq m, \tag{4-10}$$

$a_0|_{x_1=0} = \psi,\ a_k|_{x_1=0} = 0$ for $k \geq 1$, where $a_{-1} = 0$. Let us see that the transport equations have smooth solutions for $(x', \xi') \in \mathcal{U},\ 0 \leq x_1 \leq 3\delta_1$, provided δ_1 and $\mathcal{U}$ are taken small enough. As above, we will be looking for a_k in the form

$$a_k = \sum_{j=0}^{M-1} (i\theta)^j a_{k,j}(x, \xi').$$

We let $a_{k,j}$ satisfy the equations

$$2i\sum_{\nu=0}^{j} \partial_{x_1}\varphi_\nu\, \partial_{x_1}a_{k,j-\nu} + 2i\sum_{\nu=0}^{j}\langle B(x)\nabla_{x'}\varphi_\nu, \nabla_{x'}a_{k,j-\nu}\rangle + i(\Delta_X\varphi_j)a_{k,j} + \Delta_X a_{k-1,j} = 0, \tag{4-11}$$

$0 \leq j \leq M-1,\ a_{0,0}|_{x_1=0} = \psi,\ a_{k,j}|_{x_1=0} = 0$ for $k + j \geq 1$. Then the functions a_k satisfy (4-10) with $\mathcal{Q}_M^{(k)}$ polynomial in θ. As in the case of (4-5) one can solve (4-11) locally. Then we can write

$$V_- := h^{-1}\big(h^2\Delta_X + (1 + i\theta)n(x)\big)\tilde{u}_\psi^- = \mathcal{K}_1^- f + \mathcal{K}_2^- f,$$

where

$$\begin{aligned}\mathcal{K}_1^- f &= h[\Delta_X, \phi]\mathcal{K}^- f = h\big(2\phi'(x_1)\partial_{x_1} + c(x)\phi''(x_1)\big)\mathcal{K}^- f\\ &= (2\pi h)^{-d+1}\iint e^{(i/h)(\langle y', \xi'\rangle + \varphi(x, \xi', \theta))}A_1^-(x, \xi', \lambda) f(y')\, d\xi'\, dy',\end{aligned}$$

c being some smooth function and

$$A_1^- = 2i\phi' a\, \partial_{x_1}\varphi + hc\phi''\, \partial_{x_1}a,$$

and

$$(\mathcal{K}_2^- f)(x) = (2\pi h)^{-d+1} \iint e^{(i/h)(\langle y',\xi'\rangle+\varphi(x,\xi',\theta))} A_2^-(x,\xi',\lambda) f(y')\, d\xi'\, dy',$$

where

$$A_2^- = \phi(x_1)\Big(-h^{-1}\theta^M \mathcal{R}_M a + \theta^M \sum_{k=0}^{m} h^k \mathcal{Q}_M^{(k)} + h^{m+1}\Delta_X a_m\Big).$$

We claim that Proposition 4.1 follows from:

Lemma 4.2. *The function V_- satisfies the estimate*

$$\|V_-\|_{H_h^1(X)} \lesssim e^{-C|\operatorname{Im}\lambda|}\|f\| + \mathcal{O}_m(h^{m-d})\|f\| + \mathcal{O}_M(h^{\epsilon M-d})\|f\| \tag{4-12}$$

with some constant $C > 0$.

Indeed, if u_ψ^- denotes the solution to (1-1) with f replaced by $\mathrm{Op}_h(\psi) f$ and $\tilde{u}_\psi^-$ is the parametrix built above, then the function $v = u_\psi^- - \tilde{u}_\psi^-$ satisfies (3-1) with $f \equiv 0$. Therefore, by the estimates (3-10) and (4-12) we have

$$\|h\mathcal{N}(\lambda; n)\mathrm{Op}_h(\psi) - T_\psi^-\|_{L^2(\partial X)\to H_h^1(\partial X)} \lesssim e^{-C|\operatorname{Im}\lambda|} + \mathcal{O}_m(h^{m-d}) + \mathcal{O}_M(h^{\epsilon M-d}), \tag{4-13}$$

where the operator T_ψ^- is defined by

$$T_\psi^- f = h\,\partial_{x_1}\mathcal{K}^- f|_{x_1=0}.$$

Hence, in view of (4-3),

$$\begin{aligned}(T_\psi^- f)(x') &= (2\pi h)^{-d+1} \iint e^{(i/h)\langle y'-x',\xi'\rangle}\big(i\psi\,\partial_{x_1}\varphi(0,x',\xi',\theta) + h\,\partial_{x_1} a(0,x',\xi',\lambda)\big) f(y')\, d\xi'\, dy' \\ &= \mathrm{Op}_h(\rho\psi + \mathcal{O}(\theta^{M/2})) f + \sum_{k=0}^{m} h^{k+1}\mathrm{Op}_h(\partial_{x_1} a_k(0,x',\xi',\theta)) f.\end{aligned}$$

Since

$$\mathrm{Op}_h(\partial_{x_1} a_k(0,x',\xi',\theta)) = \mathcal{O}(1) : L^2(\partial X) \to H_h^1(\partial X)$$

uniformly in θ, it follows from (4-13) that

$$\big\|h\mathcal{N}(\lambda; n)\mathrm{Op}_h(\psi) - \mathrm{Op}_h(\rho\psi)\big\|_{L^2(\partial X)\to H_h^1(\partial X)} \lesssim e^{-C|\operatorname{Im}\lambda|} + \mathcal{O}(h). \tag{4-14}$$

On the other hand, using a suitable partition of the unity we can write the function χ_δ^- as $\sum_{j=1}^{J}\psi_j$, where each function ψ_j has the same properties as the function ψ above. In other words, we have (4-14) with ψ replaced by each ψ_j, which after summing up leads to (4-1). □

Proof of Lemma 4.2. Let α be a multi-index such that $|\alpha| \le 1$. Since

$$i|\alpha| A_2^- \partial_x^\alpha \varphi + (h\,\partial_x)^\alpha A_2^- = \mathcal{O}_m(h^{m+1}) + \mathcal{O}_M(h^{\epsilon M-1})$$

and $\operatorname{Im}\varphi \ge 0$, the kernel of the operator $(h\,\partial_x)^\alpha \mathcal{K}_2^- : L^2(\partial X) \to L^2(X)$ is $\mathcal{O}_m(h^{m-d}) + \mathcal{O}_M(h^{\epsilon M-d})$, and hence so is its norm. Since the function A_1^- is supported in the interval $[\delta_1/2, 3\delta_1]$ with respect to the variable x_1, to bound the norm of the operator $\mathcal{K}^-_{1,\alpha} := (h\,\partial_x)^\alpha \mathcal{K}^-_1 : L^2(\partial X) \to L^2(X)$ it suffices to show that

$$\|\mathcal{K}^-_{1,\alpha}\|_{L^2(\partial X)\to L^2(\partial X)} \lesssim e^{-C|\theta|/h} + \mathcal{O}(h^\infty) \tag{4-15}$$

uniformly in $x_1 \in [\delta_1/2, 3\delta_1]$. Since $|\theta|/h \sim |\operatorname{Im}\lambda|$, (4-15) will imply (4-12). We would like to consider $\mathcal{K}^-_{1,\alpha}$ as an h-FIO with phase $\operatorname{Re}\varphi$ and amplitude

$$A_\alpha = e^{-\operatorname{Im}\varphi/h}(i|\alpha| A^-_1 \partial^\alpha_x \varphi + (h\,\partial_x)^\alpha A^-_1).$$

To do so, we need to have that the phase satisfies the condition

$$\left|\det\left(\frac{\partial^2 \operatorname{Re}\varphi}{\partial x' \partial \xi'}\right)\right| \geq \widetilde{C} > 0 \tag{4-16}$$

for $|\theta|$ small enough, where $\widetilde{C}$ is a constant independent of θ. Since $\operatorname{Re}\varphi = \varphi_0 + \mathcal{O}(|\theta|)$, it suffices to show (4-16) for the phase φ_0. This, however, is easy to arrange by taking x_1 small enough because $\varphi_0 = -\langle x', \xi'\rangle + \mathcal{O}(x_1)$ and (4-16) is trivially fulfilled for the phase $-\langle x', \xi'\rangle$. On the other hand, using that $\operatorname{Im}\varphi = \mathcal{O}(|\theta|)$ together with (4-9) we get the following bounds for the amplitude:

$$|\partial^{\beta_1}_{x'} \partial^{\beta_2}_{\xi'} A_\alpha| \leq C_{\beta_1,\beta_2} \sum_{0\leq k\leq |\beta_1|+|\beta_2|} \left(\frac{|\theta|}{h}\right)^k e^{-C\delta_1|\theta|/(8h)} \leq \widetilde{C}_{\beta_1,\beta_2} e^{-C\delta_1|\theta|/(9h)} \tag{4-17}$$

for all multi-indices β_1 and β_2. It follows from (4-16) and (4-17) that, mod $\mathcal{O}(h^\infty)$, the operator $(\mathcal{K}^-_{1,\alpha})^* \mathcal{K}^-_{1,\alpha}$ is an h-ΨDO in the class $\mathrm{OP}S^0_0(\partial X)$ uniformly in θ with a symbol which is $\mathcal{O}(e^{-2C|\theta|/h})$ together with all derivatives, where $C > 0$ is a new constant. Therefore, its norm is also $\mathcal{O}(e^{-2C|\theta|/h})$, which clearly implies (4-15). □

5. Parametrix construction in the elliptic region

We keep the notations from the previous sections and note that $\rho\chi^+_\delta \in S^1_0(\partial X)$. It is easy also to see that $0 < C_1\langle\xi'\rangle \leq |\rho| \leq C_2\langle\xi'\rangle$ on $\operatorname{supp}\chi^+_\delta$, where C_1 and C_2 are constants depending on δ. In this section we will prove:

Proposition 5.1. *There is a constant $C > 0$ depending on δ but independent of λ such that*

$$\left\| h\mathcal{N}(\lambda; n)\operatorname{Op}_h(\chi^+_\delta) - \operatorname{Op}_h(\rho\chi^+_\delta + hb) \right\|_{L^2(\partial X)\to H^1_h(\partial X)} \leq Ch, \tag{5-1}$$

where $b \in S^0_0(\partial X)$ does not depend on λ or the function n.

Proof. The estimate (5-1) is a consequence of the parametrix built in [Vodev 2015]. In what follows we will recall this construction. We will first proceed locally and then we will use partition of the unity to get the global parametrix. Fix a point $x^0 \in \partial X$ and let $\mathcal{U}_0 \subset \partial X$ be a small open neighbourhood of x^0. Let (x_1, x'), $x_1 > 0$, $x' \in \mathcal{U}_0$, be the normal coordinates used in the previous section. Take a function $\psi^0 \in C^\infty_0(\mathcal{U}_0)$ and set $\psi = \psi^0\chi^+_\delta$. As in the previous section, we will construct a parametrix $\tilde{u}^+_\psi$ of the solution of (1-1) with $\tilde{u}^+_\psi|_{x_1=0} = \operatorname{Op}_h(\psi) f$ in the form $\tilde{u}^+_\psi = \phi(x_1)\mathcal{K}^+ f$, where $\phi(x_1) = \chi(x_1/\delta_1)$, $0 < \delta_1 \ll 1$, is a parameter independent of λ to be fixed later on, and

$$(\mathcal{K}^+ f)(x) = (2\pi h)^{-d+1} \iint e^{(i/h)(\langle y', \xi'\rangle + \varphi(x,\xi',\theta))} a(x, \xi', \lambda) f(y')\, d\xi'\, dy'.$$

The phase φ is complex-valued such that $\varphi|_{x_1=0} = -\langle x', \xi'\rangle$ and satisfies the eikonal equation mod $\mathcal{O}(x_1^M)$:

$$(\partial_{x_1}\varphi)^2 + \langle B(x)\nabla_{x'}\varphi, \nabla_{x'}\varphi\rangle - (1+i\theta)n(x) = x_1^M \widetilde{\mathcal{R}}_M, \tag{5-2}$$

where $M \gg 1$ is an arbitrary integer, and the function $\widetilde{\mathcal{R}}_M$ is smooth up to the boundary $x_1 = 0$. It is shown in [Vodev 2015, Section 4] that for $(x', \xi') \in \operatorname{supp}\psi$, (5-2) has a smooth solution of the form

$$\varphi = \sum_{k=0}^{M-1} x_1^k \varphi_k(x', \xi', \theta), \quad \varphi_0 = -\langle x', \xi'\rangle,$$

satisfying

$$\partial_{x_1}\varphi|_{x_1=0} = \varphi_1 = -i\rho. \tag{5-3}$$

Moreover, taking δ_1 small enough we can arrange that

$$\operatorname{Im}\varphi \geq -\tfrac{1}{2}x_1 \operatorname{Re}\rho \geq C x_1 \langle \xi'\rangle, \quad C > 0, \tag{5-4}$$

for $0 \leq x_1 \leq 3\delta_1$, $(x', \xi') \in \operatorname{supp}\psi$. The amplitude a is of the form

$$a = \sum_{j=0}^{m} h^j a_j(x, \xi', \theta),$$

where $m \gg 1$ is an arbitrary integer and the functions a_j satisfy the transport equations mod $\mathcal{O}(x_1^M)$:

$$2i\partial_{x_1}\varphi\partial_{x_1}a_j + 2i\langle B(x)\nabla_{x'}\varphi, \nabla_{x'}a_j\rangle + i(\Delta_X\varphi)a_j + \Delta_X a_{j-1} = x_1^M \widetilde{\mathcal{Q}}_M^{(j)}, \quad 0 \leq j \leq m, \tag{5-5}$$

$a_0|_{x_1=0} = \psi$, $a_j|_{x_1=0} = 0$ for $j \geq 1$, where $a_{-1} = 0$ and the functions $\widetilde{\mathcal{Q}}_M^{(j)}$ are smooth up to the boundary $x_1 = 0$. It is shown in [Vodev 2015, Section 4] that the equations (5-5) have unique smooth solutions of the form

$$a_j = \sum_{k=0}^{M-1} x_1^k a_{k,j}(x', \xi', \theta)$$

with functions $a_{k,j} \in S_0^{-j}(\partial X)$ uniformly in θ. We can write

$$V_+ := h^{-1}\big(h^2\Delta_X + (1+i\theta)n(x)\big)\tilde{u}_\psi^+ = \mathcal{K}_1^+ f + \mathcal{K}_2^+ f,$$

where

$$\begin{aligned}\mathcal{K}_1^+ f &= h[\Delta_X, \phi]\mathcal{K}^+ f = h\big(2\phi'(x_1)\partial_{x_1} + c(x)\phi''(x_1)\big)\mathcal{K}^+ f \\ &= (2\pi h)^{-d+1} \iint e^{(i/h)(\langle y', \xi'\rangle + \varphi(x, \xi', \theta))} A_1^+(x, \xi', \lambda) f(y')\, d\xi'\, dy',\end{aligned}$$

with

$$A_1^+ = 2i\phi' a \partial_{x_1}\varphi + hc\phi''\partial_{x_1} a,$$

and

$$(\mathcal{K}_2^+ f)(x) = (2\pi h)^{-d+1} \iint e^{(i/h)(\langle y', \xi'\rangle + \varphi(x, \xi', \theta))} A_2^+(x, \xi', \lambda) f(y')\, d\xi'\, dy',$$

where

$$A_2^+ = \phi(x_1)\bigg(-h^{-1}x_1^M \widetilde{\mathcal{R}}_M a + x_1^M \sum_{j=0}^{m} h^j \widetilde{\mathcal{Q}}_M^{(j)} + h^{m+1}\Delta_X a_m\bigg).$$

As in the previous section, we will derive Proposition 5.1 from (5-3) and the following:

Lemma 5.2. *The function V_+ satisfies the estimate*

$$\|V_+\|_{H_h^1(X)} \le \mathcal{O}_m(h^{m-d})\|f\| + \mathcal{O}_M(h^{M-d})\|f\|. \tag{5-6}$$

Proof. Let α be a multi-index such that $|\alpha| \le 1$. In view of (5-4) we have

$$\left|e^{i\varphi/h}(i|\alpha|A_1^+\partial_x^\alpha\varphi + (h\,\partial_x)^\alpha A_1^+)\right| \lesssim \sup_{\delta_1/2\le x_1\le 3\delta_1} e^{-\operatorname{Im}\varphi/h} \lesssim e^{-C\langle\xi'\rangle/h} = \mathcal{O}_M((h/\langle\xi'\rangle)^M)$$

for every integer $M \gg 1$. Therefore, the kernel of the operator $(h\,\partial_x)^\alpha\mathcal{K}_1^+ : L^2(\partial X) \to L^2(X)$ is $\mathcal{O}_M(h^{M-d+1})$, and hence so is its norm. By (5-4) we also have

$$x_1^M e^{-\operatorname{Im}\varphi/h} \le x_1^M e^{-Cx_1\langle\xi'\rangle/h} = \mathcal{O}_M((h/\langle\xi'\rangle)^M).$$

This implies

$$e^{i\varphi/h}\big(i|\alpha|A_2^+\partial_x^\alpha\varphi + (h\,\partial_x)^\alpha A_2^+\big) = \mathcal{O}_M((h/\langle\xi'\rangle)^{M-1}) + \mathcal{O}_m((h/\langle\xi'\rangle)^m),$$

which again implies the desired bound for the norm of the operator $(h\,\partial_x)^\alpha\mathcal{K}_2^+$. □

By the estimates (3-10) and (5-6) we have

$$\left\|h\mathcal{N}(\lambda;n)\mathrm{Op}_h(\psi) - T_\psi^+\right\|_{L^2(\partial X)\to H_h^1(\partial X)} \le \mathcal{O}_m(h^{m-d}) + \mathcal{O}_M(h^{M-d}), \tag{5-7}$$

where the operator T_ψ^+ is defined by

$$T_\psi^+ f = h\,\partial_{x_1}\mathcal{K}^+ f|_{x_1=0}.$$

In view of (5-3), we have

$$\begin{aligned}(T_\psi^+ f)(x') &= (2\pi h)^{-d+1}\iint e^{(i/h)\langle y'-x',\xi'\rangle}\big(i\psi\,\partial_{x_1}\varphi(0,x',\xi',\theta) + h\,\partial_{x_1}a(0,x',\xi',\lambda)\big)f(y')\,d\xi'\,dy' \\ &= \mathrm{Op}_h(\rho\psi)f + \sum_{j=0}^m h^{j+1}\mathrm{Op}_h(a_{1,j}(x',\xi',\theta))f,\end{aligned}$$

where $a_{1,j} \in S_0^{-j}(\partial X)$. Hence

$$\mathrm{Op}_h(a_{1,j}) = \mathcal{O}(1) : L^2(\partial X) \to H_h^j(\partial X).$$

Therefore it follows from (5-7) that

$$\left\|h\mathcal{N}(\lambda;n)\mathrm{Op}_h(\psi) - \mathrm{Op}_h(\rho\psi + ha_{1,0})\right\|_{L^2(\partial X)\to H_h^1(\partial X)} \le \mathcal{O}(h). \tag{5-8}$$

We need now the following:

Lemma 5.3. *There exists a function $b^0 \in S_0^0(\partial X)$, independent of λ and n, such that*

$$a_{1,0} - b^0 \in S_0^{-1}(\partial X). \tag{5-9}$$

Proof. We will calculate the function $a_{1,0}$ explicitly. Note that this lemma (as well as Proposition 5.1) is also used in [Vodev 2015], but the proof therein is not correct since $a_{1,0}$ is calculated incorrectly. Therefore we will give here a new proof. Clearly, it suffices to prove (5-9) with $a_{1,0}$ replaced by $(1-\eta)a_{1,0}$ with some function $\eta \in C_0^\infty(T^*\partial X)$ independent of h. Since $\rho = -\sqrt{r_0}(1+\mathcal{O}(r_0^{-1}))$ as $r_0 \to \infty$, it is easy to see that

$$(1-\eta)\rho^{-k} - (1-\eta)(-\sqrt{r_0})^{-k} \in S_0^{-k-1}(\partial X) \tag{5-10}$$

for every integer $k \geq 0$, provided η is taken such that $\eta = 1$ for $|\xi'| \leq A$ with some $A > 1$ big enough. We will now calculate the function φ_2 from the eikonal equation. To this end, write

$$B(x) = B_0(x') + x_1 B_1(x') + \mathcal{O}(x_1^2), \quad n(x) = n_0(x') + x_1 n_1(x') + \mathcal{O}(x_1^2)$$

and observe that the left-hand side of (5-2) is equal to

$$x_1\big(4\varphi_1\varphi_2 + 2\langle B_0\nabla_{x'}\varphi_0, \nabla_{x'}\varphi_1\rangle + \langle B_1\nabla_{x'}\varphi_0, \nabla_{x'}\varphi_0\rangle - (1+i\theta)n_1\big) + \mathcal{O}(x_1^2).$$

Hence, taking into account that $\varphi_0 = -\langle x', \xi'\rangle$ and $\varphi_1 = -i\rho$, we get

$$\varphi_2 = (2\rho)^{-1}\langle B_0\xi', \nabla_{x'}\rho\rangle + (4i\rho)^{-1}\langle B_1\xi', \xi'\rangle - (1+i\theta)(4i\rho)^{-1}n_1.$$

Using the identity

$$2\rho\nabla_{x'}\rho = \nabla_{x'}r_0 - (1+i\theta)\nabla_{x'}n_0$$

we can write φ_2 in the form

$$\varphi_2 = (2\rho)^{-2}\langle B_0\xi', \nabla_{x'}r_0\rangle + (4i\rho)^{-1}\langle B_1\xi', \xi'\rangle - (1+i\theta)(2\rho)^{-2}\langle B_0\xi', \nabla_{x'}n_0\rangle - (1+i\theta)(4i\rho)^{-1}n_1.$$

By (5-10) we conclude that, mod $S_0^{-1}(\partial X)$,

$$(1-\eta)\frac{\varphi_2}{\varphi_1} = -i4^{-1}(1-\eta)r_0^{-3/2}\langle B_0\xi', \nabla_{x'}r_0\rangle + (1-\eta)(4r_0)^{-1}\langle B_1\xi', \xi'\rangle. \tag{5-11}$$

Write now the operator Δ_X in the form

$$\Delta_X = \partial_{x_1}^2 + \langle B_0\nabla_{x'}, \nabla_{x'}\rangle + q_1(x')\partial_{x_1} + \langle q_2(x'), \nabla_{x'}\rangle + \mathcal{O}(x_1)$$

and observe that

$$\Delta_X\varphi = 2\varphi_2 + q_1\varphi_1 - \langle q_2(x'), \xi'\rangle + \mathcal{O}(x_1).$$

We now calculate the left-hand side of (5-5) with $j = 0$ modulo $\mathcal{O}(x_1)$. Recall that $a_{0,0} = \psi$. We obtain

$$2i\varphi_1 a_{1,0} + 2i\langle B_0\nabla_{x'}\varphi_0, \nabla_{x'}a_{0,0}\rangle + i(\Delta_X\varphi)a_{0,0} = 2i\varphi_1 a_{1,0} + 2i\langle B_0\xi', \nabla_{x'}\psi\rangle + i\big(2\varphi_2 + q_1\varphi_1 - \langle q_2(x'), \xi'\rangle\big)\psi.$$

Since the right-hand side is $\mathcal{O}(x_1^M)$, the above function must be identically zero. Thus we get the following expression for the function $a_{1,0}$:

$$a_{1,0} = -\varphi_1^{-1}\langle B_0\xi', \nabla_{x'}\psi\rangle - \big(\varphi_1^{-1}\varphi_2 + 2^{-1}q_1 - (2\varphi_1)^{-1}\langle q_2(x'), \xi'\rangle\big)\psi. \tag{5-12}$$

Taking into account that $\psi = \psi^0$ on $\mathrm{supp}(1-\eta)$, we find from (5-10)–(5-12) that (5-9) holds with

$$b^0 = i(1-\eta)r_0^{-1/2}\langle B_0\xi', \nabla_{x'}\psi^0\rangle$$
$$-4^{-1}(1-\eta)\psi^0\big(-ir_0^{-3/2}\langle B_0\xi', \nabla_{x'}r_0\rangle + r_0^{-1}\langle B_1\xi', \xi'\rangle + 2q_1 + 2r_0^{-1/2}\langle q_2(x'), \xi'\rangle\big). \quad (5\text{-}13)$$

Clearly, $b^0 \in S_0^0(\partial X)$ is independent of λ and n, as desired. □

Lemma 5.3 implies that

$$\mathrm{Op}_h(a_{1,0} - b^0) = \mathcal{O}(1) : L^2(\partial X) \to H_h^1(\partial X). \quad (5\text{-}14)$$

Now, using a suitable partition of the unity on ∂X we can write $1 = \sum_{j=1}^J \psi_j^0$. Hence, we can write the function χ_δ^+ as $\sum_{j=1}^J \psi_j$, where $\psi_j = \psi_j^0\chi_\delta^+$. Since we have (5-8) and (5-14) with ψ replaced by each ψ_j, we get (5-1) by summing up all the estimates. □

It follows from the estimate (3-11) applied with $V \equiv 0$ that

$$h\mathcal{N}(\lambda; n)\mathrm{Op}_h(\chi_\delta^0) = \mathcal{O}(\delta) : L^2(\partial X) \to H_h^1(\partial X) \quad (5\text{-}15)$$

provided $|\mathrm{Im}\,\lambda| \ge \delta^{-4}$ and $\mathrm{Re}\,\lambda \ge C_\delta \gg 1$. Now Theorem 1.2 follows from (5-15) and Propositions 4.1 and 5.1. Let us now see that Theorem 1.1 follows from Theorem 1.2. Since the operator $-h^2\Delta_{\partial X} \ge 0$ is self-adjoint, we have the bound

$$\begin{aligned}\big\|hp(-\Delta_{\partial X})\chi_2((-h^2\Delta_{\partial X}-1)\delta^{-2})\big\| &= \big\|\sqrt{-h^2\Delta_{\partial X}-1-i\theta}\,\chi((-h^2\Delta_{\partial X}-1)\delta^{-2})\big\| \\ &\le \sup_{\sigma\ge 0}\big|\sqrt{\sigma-1-i\theta}\,\chi((\sigma-1)\delta^{-2})\big| \\ &\le \sup_{\delta^2\le|\sigma-1|\le 2\delta^2}\sqrt{|\sigma-1|+|\theta|} \le \mathcal{O}(\delta+|\theta|^{1/2}) = \mathcal{O}(\delta+h^{\epsilon/2}). \quad (5\text{-}16)\end{aligned}$$

On the other hand, it is well known that the operator $hp(-\Delta_{\partial X})(1-\chi_2)((-h^2\Delta_{\partial X}-1)\delta^{-2})$ is an h-ΨDO in the class $\mathrm{OP}S_0^1(\partial X)$ with principal symbol $\rho(1-\chi_\delta^0)$. This implies the bound

$$hp(-\Delta_{\partial X})(1-\chi_2)((-h^2\Delta_{\partial X}-1)\delta^{-2}) - \mathrm{Op}_h(\rho(1-\chi_\delta^0)) = \mathcal{O}(h) : L^2(\partial X) \to L^2(\partial X). \quad (5\text{-}17)$$

It is easy to see that Theorem 1.1 follows from (1-3) together with (5-16) and (5-17). □

6. Proof of Theorem 2.1

Define the DN maps $\mathcal{N}_j(\lambda)$, $j = 1, 2$, by

$$\mathcal{N}_j(\lambda) f = \partial_\nu u_j|_\Gamma,$$

where ν is the Euclidean unit normal to Γ and u_j is the solution to the equation

$$\begin{cases}(\nabla c_j(x)\nabla + \lambda^2 n_j(x))u_j = 0 & \text{in } \Omega,\\ u_j = f & \text{on } \Gamma,\end{cases} \quad (6\text{-}1)$$

and consider the operator

$$T(\lambda) = c_1\mathcal{N}_1(\lambda) - c_2\mathcal{N}_2(\lambda).$$

Clearly, λ is a transmission eigenvalue if there exists a nontrivial function f such that $T(\lambda)f = 0$. Therefore Theorem 2.1 is a consequence of the following:

Theorem 6.1. *Under the conditions of Theorem 2.1, the operator* $T(\lambda)$ *sends* $H^{(1+k)/2}(\Gamma)$ *into* $H^{(1-k)/2}(\Gamma)$, *where* $k=-1$ *if* (2-2) *holds and* $k=1$ *if* (2-4) *holds. Moreover, there exists a constant* $C>0$ *such that* $T(\lambda)$ *is invertible for* $\operatorname{Re}\lambda \geq 1$ *and* $|\operatorname{Im}\lambda| \geq C$ *with an inverse satisfying in this region the bound*

$$\|T(\lambda)^{-1}\|_{H^{(1-k)/2}(\Gamma)\to H^{(1+k)/2}(\Gamma)} \lesssim |\lambda|^{(k-1)/2}, \tag{6-2}$$

where the Sobolev spaces are equipped with the classical norms.

Proof. We may suppose that $\lambda \in \Lambda_\epsilon = \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda \geq C_\epsilon \gg 1,\ |\operatorname{Im}\lambda| \leq |\lambda|^\epsilon\},\ 0<\epsilon \ll 1$, since the case when $\lambda \in \{\operatorname{Re}\lambda \geq 1\} \setminus \Lambda_\epsilon$ follows from the analysis in [Vodev 2015]. We will equip the boundary Γ with the Riemannian metric induced by the Euclidean metric g_E in Ω and will denote by r_0 the principal symbol of the Laplace–Beltrami operator $-\Delta_\Gamma$. We would like to apply Theorem 1.2 to the operators $\mathcal{N}_j(\lambda)$. However, some modifications must be done coming from the presence of the function c_j in (6-1). Indeed, in the definition of the operator $\mathcal{N}(\lambda; n)$ in Section 1, the normal derivative is taken with respect to the Riemannian metric $g_j = c_j^{-1} g_E$, while in the definition of the operator $\mathcal{N}_j(\lambda)$ it is taken with respect to the metric g_E. The first observation to be done is that the glancing region corresponding to the problem (6-1) is defined by $\Sigma_j := \{(x',\xi') \in T^*\Gamma : r_j(x',\xi') = 1\}$, where $r_j := m_j^{-1} r_0,\ m_j := (n_j/c_j)|_\Gamma$. We define now the cut-off functions $\chi^0_{\delta,j}$ by replacing in the definition of χ^0_δ the function $r_\sharp$ by r_j. Secondly, the function ρ must be replaced by

$$\rho_j(x',\xi') = \sqrt{r_0(x',\xi') - (1+i\theta)m_j(x')}, \quad \operatorname{Re}\rho_j < 0.$$

With these changes, the operator $\mathcal{N}_j(\lambda)$ satisfies the estimate (1-3). Set

$$\tau_\delta = c_1\rho_1(1-\chi^0_{\delta,1}) - c_2\rho_2(1-\chi^0_{\delta,2}) = \tau - c_1\rho_1\chi^0_{\delta,1} + c_2\rho_2\chi^0_{\delta,2},$$

where

$$\tau = c_1\rho_1 - c_2\rho_2 = \frac{\tilde{c}(x')\big(c_0(x')r_0(x',\xi') - 1 - i\theta\big)}{c_1\rho_1 + c_2\rho_2}, \tag{6-3}$$

where $\tilde{c}$ and c_0 are the restrictions on Γ of the functions

$$c_1n_1 - c_2n_2 \quad \text{and} \quad \frac{c_1^2 - c_2^2}{c_1n_1 - c_2n_2}$$

respectively. Clearly, under the conditions of Theorem 2.1, we have $\tilde{c}(x') \neq 0$ for all $x' \in \Gamma$. Moreover, (2-2) implies $c_0 \equiv 0$, while (2-4) implies $c_0(x') < 0$ for all $x' \in \Gamma$. Hence,

$$0 < C_1 \leq |c_0r_0 - 1 - i\theta| \leq C_2$$

if (2-2) holds, and

$$0 < C_1\langle r_0\rangle \leq |c_0r_0 - 1 - i\theta| \leq C_2\langle r_0\rangle$$

if (2-4) holds. Using this, together with (6-3), and the fact that $\rho_j \sim -\sqrt{r_0}$ as $r_0 \to \infty$, we get

$$0 < C_1'\langle\xi'\rangle^k \leq C_1\langle r_0\rangle^{k/2} \leq |\tau| \leq C_2\langle r_0\rangle^{k/2} \leq C_2'\langle\xi'\rangle^k, \tag{6-4}$$

where $k=-1$ if (2-2) holds and $k=1$ if (2-4) holds. Let $\eta\in C_0^\infty(T^*\Gamma)$ be such that $\eta=1$ on $|\xi'|\le A$ and $\eta=0$ on $|\xi'|\ge A+1$, where $A\gg 1$ is a big parameter independent of λ and δ. Taking A big enough we can arrange that $(1-\eta)\tau_\delta=(1-\eta)\tau$. On the other hand, we have $\eta\tau_\delta=\eta\tau+\mathcal{O}(\delta+|\theta|^{1/2})$. Therefore, taking δ and $|\theta|$ small enough, we get from (6-4) that the function τ_δ satisfies the bounds

$$\widetilde{C}_1\langle\xi'\rangle^k\le|\tau_\delta|\le\widetilde{C}_2\langle\xi'\rangle^k \tag{6-5}$$

with positive constants $\widetilde{C}_1$ and $\widetilde{C}_2$ independent of δ and θ. Furthermore, one can easily check that $(1-\eta)\tau\in S_0^k(\Gamma)$ and $\eta\tau_\delta\in S_0^{-2}(\Gamma)$. Hence, $\tau_\delta\in S_0^k(\Gamma)$, which in turn implies that the operator $\mathrm{Op}_h(\tau_\delta)$ sends $H^{(1+k)/2}(\Gamma)$ into $H^{(1-k)/2}(\Gamma)$. Moreover, it follows from (6-5) that the operator $\mathrm{Op}_h(\tau_\delta):H_h^{(1+k)/2}(\Gamma)\to H_h^{(1-k)/2}(\Gamma)$ is invertible with an inverse satisfying the bound

$$\|\mathrm{Op}_h(\tau_\delta)^{-1}\|_{H_h^{(1-k)/2}(\Gamma)\to H_h^{(1+k)/2}(\Gamma)}\le\widetilde{C} \tag{6-6}$$

with a constant $\widetilde{C}>0$ independent of λ and δ. We now apply Theorem 2.1 to the operators $\mathcal{N}_j(\lambda)$. We get, for $\lambda\in\Lambda_\epsilon,\ |\mathrm{Im}\,\lambda|\ge C_\delta\gg 1,\ \mathrm{Re}\,\lambda\ge C_{\epsilon,\delta}\gg 1$, that

$$\|hT(\lambda)-\mathrm{Op}_h(\tau_\delta)\|_{L^2(\Gamma)\to L^2(\Gamma)}\le C\delta \tag{6-7}$$

in the anisotropic case, and

$$\|hT(\lambda)-\mathrm{Op}_h(\tau_\delta)\|_{L^2(\Gamma)\to H_h^1(\Gamma)}\le C\delta \tag{6-8}$$

in the isotropic case, where $C>0$ is a constant independent of λ and δ. We introduce the operators

$$\mathcal{A}_1(\lambda)=(hT(\lambda)-\mathrm{Op}_h(\tau_\delta))\mathrm{Op}_h(\tau_\delta)^{-1},$$
$$\mathcal{A}_2(\lambda)=\mathrm{Op}_h(\tau_\delta)^{-1}(hT(\lambda)-\mathrm{Op}_h(\tau_\delta)).$$

It follows from (6-6)–(6-8) that in the anisotropic case we have the bound

$$\|\mathcal{A}_1(\lambda)\|_{L^2(\Gamma)\to L^2(\Gamma)}\le C'\delta, \tag{6-9}$$

while in the isotropic case we have the bound

$$\|\mathcal{A}_2(\lambda)\|_{L^2(\Gamma)\to L^2(\Gamma)}\le C'\delta, \tag{6-10}$$

where $C'>0$ is a constant independent of λ and δ. Hence, taking δ small enough we can arrange that the operators $1+\mathcal{A}_j(\lambda)$ are invertible on $L^2(\Gamma)$ with inverses whose norms are bounded by 2. We now write the operator $hT(\lambda)$ as

$$hT(\lambda)=(1+\mathcal{A}_1(\lambda))\mathrm{Op}_h(\tau_\delta)$$

in the anisotropic case, and as

$$hT(\lambda)=\mathrm{Op}_h(\tau_\delta)(1+\mathcal{A}_2(\lambda))$$

in the isotropic case. Therefore, the operator $hT(\lambda)$ is invertible in the desired region and by (6-6) we get the bound

$$\|(hT(\lambda))^{-1}\|_{H_h^{(1-k)/2}(\Gamma)\to H_h^{(1+k)/2}(\Gamma)}\le 2\widetilde{C}. \tag{6-11}$$

Passing from semiclassical to classical Sobolev norms, one can easily see that (6-11) implies (6-2). □

7. Proof of Theorem 2.2

We keep the notations from the previous section. Theorem 2.2 is a consequence of the following:

Theorem 7.1. *Under the conditions of Theorem 2.2, there exists a constant $C > 0$ such that the operator $T(\lambda) : H^1(\Gamma) \to L^2(\Gamma)$ is invertible for $\operatorname{Re}\lambda \geq 1$ and $|\operatorname{Im}\lambda| \geq C \log(\operatorname{Re}\lambda + 1)$ with an inverse satisfying in this region the bound*

$$\|T(\lambda)^{-1}\|_{L^2(\Gamma)\to L^2(\Gamma)} \lesssim 1. \tag{7-1}$$

Proof. As in the previous section we may suppose that $\lambda \in \Lambda_\epsilon$. We will again make use of the identity (6-3) with the difference that under the condition (2-6) we have $c_0(x') > 0$ for all $x' \in \Gamma$. This means that $|\tau|$ can get small near the characteristic variety $\Sigma = \{(x', \xi') \in T^*\Gamma : r(x', \xi') = 1\}$, where $r := c_0 r_0$. Clearly, the assumption (2-7) implies that $\Sigma_1 \cap \Sigma_2 = \varnothing$. This in turn implies that $\Sigma \cap \Sigma_j = \varnothing,\ j = 1, 2$. Indeed, if we suppose that there is a $\zeta^0 \in \Sigma \cap \Sigma_j$ for $j = 1$ or $j = 2$, then it is easy to see that $\zeta^0 \in \Sigma_1 \cap \Sigma_2$, which however is impossible in view of (2-7). Therefore, we can choose a cut-off function $\chi^0 \in C^\infty(T^*\Gamma)$ such that $\chi^0 = 1$ in a small neighbourhood of Σ, $\chi^0 = 0$ outside another small neighbourhood of Σ, and $\operatorname{supp}\chi^0 \cap \Sigma_j = \varnothing,\ j = 1, 2$. This means that $\operatorname{supp}\chi^0$ belongs either to the hyperbolic region $\{r_j \leq 1 - \delta^2\}$ or to the elliptic region $\{r_j \geq 1 + \delta^2\}$, provided $\delta > 0$ is taken small enough. Therefore, we can use Propositions 4.1 and 5.1 to get the estimate

$$\left\| h\mathcal{N}_j(\lambda)\operatorname{Op}_h(\chi^0) - \operatorname{Op}_h(\rho_j\chi^0) \right\|_{L^2(\Gamma)\to L^2(\Gamma)} \lesssim h + e^{-C|\operatorname{Im}\lambda|},$$

which implies

$$\left\| hT(\lambda)\operatorname{Op}_h(\chi^0) - \operatorname{Op}_h(\tau\chi^0) \right\|_{L^2(\Gamma)\to L^2(\Gamma)} \lesssim h + e^{-C|\operatorname{Im}\lambda|}. \tag{7-2}$$

It follows from (6-3) that near Σ the function τ is of the form $\tau = \tau_0(r - 1 - i\theta)$ with some smooth function $\tau_0 \neq 0$. We now extend τ_0 globally on $T^*\Gamma$ to a function $\tilde\tau_0 \in S_0^0(\Gamma)$ such that $\tilde\tau_0 = \tau_0$ on $\operatorname{supp}\chi^0$ and $|\tilde\tau_0| \geq \text{const.} > 0$ on $T^*\Gamma$. Hence, we can write the operator $\operatorname{Op}_h(\tau\chi^0)$ as

$$\operatorname{Op}_h(\tau\chi^0) = \operatorname{Op}_h(\chi^0)\operatorname{Op}_h(\tilde\tau_0)(\mathcal{B} - i\theta) + \mathcal{O}(h),$$

where $\mathcal{B} = \frac{1}{2}\operatorname{Op}_h(r-1) + \frac{1}{2}\operatorname{Op}_h(r-1)^*$ is a self-adjoint operator. Hence

$$(\mathcal{B} - i\theta)^{-1} = \mathcal{O}(|\theta|^{-1}) : L^2(\Gamma) \to L^2(\Gamma).$$

Since $\tilde\tau_0$ is globally elliptic, we also have

$$\operatorname{Op}_h(\tilde\tau_0)^{-1} = \mathcal{O}(1) : L^2(\Gamma) \to L^2(\Gamma).$$

This implies

$$K_1 := \operatorname{Op}_h(\chi^0)(\mathcal{B} - i\theta)^{-1}\operatorname{Op}_h(\tilde\tau_0)^{-1} = \mathcal{O}(|\theta|^{-1}) : L^2(\Gamma) \to L^2(\Gamma)$$

and (7-2) leads to the estimate

$$\|hT(\lambda)K_1 - \operatorname{Op}_h(\chi^0)\|_{L^2(\Gamma)\to L^2(\Gamma)} \lesssim |\theta|^{-1}(h + e^{-C|\operatorname{Im}\lambda|}) \lesssim |\operatorname{Im}\lambda|^{-1} + \operatorname{Re}\lambda\, e^{-C|\operatorname{Im}\lambda|} \leq \delta \tag{7-3}$$

for any $0 < \delta \ll 1$, provided $|\mathrm{Im}\,\lambda| \ge C_\delta \log(\mathrm{Re}\,\lambda)$, $\mathrm{Re}\,\lambda \ge \widetilde{C}_\delta$ with some constants $C_\delta, \widetilde{C}_\delta > 0$. On the other hand, by Theorem 1.2 we have, for $\lambda \in \Lambda_\epsilon$, $|\mathrm{Im}\,\lambda| \ge C_\delta \gg 1$, $\mathrm{Re}\,\lambda \ge C_{\epsilon,\delta} \gg 1$,

$$\left\| hT(\lambda)\mathrm{Op}_h(1-\chi^0) - \mathrm{Op}_h(\tau_\delta(1-\chi^0)) \right\|_{L^2(\Gamma)\to L^2(\Gamma)} \le C\delta. \tag{7-4}$$

As in the proof of (6-5), one can see that the function τ_δ satisfies

$$\widetilde{C}_1\langle\xi'\rangle \le |\tau_\delta| \le \widetilde{C}_2\langle\xi'\rangle \quad \text{on } \operatorname{supp}(1-\chi^0) \tag{7-5}$$

with positive constants $\widetilde{C}_1$ and $\widetilde{C}_2$ independent of δ and θ. Moreover, $\tau_\delta \in S_0^1(\Gamma)$. We extend the function τ_δ on the whole of $T^*\Gamma$ to a function $\tilde{\tau}_\delta \in S_0^1(\Gamma)$ such that $\tilde{\tau}_\delta(1-\chi^0) = \tau_\delta(1-\chi^0)$ and

$$\widetilde{C}_1'\langle\xi'\rangle \le |\tilde{\tau}_\delta| \le \widetilde{C}_2'\langle\xi'\rangle \quad \text{on } T^*\Gamma. \tag{7-6}$$

Hence

$$\|\mathrm{Op}_h(\tilde{\tau}_\delta)^{-1}\|_{L^2(\Gamma)\to L^2(\Gamma)} \le \widetilde{C} \tag{7-7}$$

with a constant $\widetilde{C} > 0$ independent of λ and δ. By (7-4) and (7-7) we obtain

$$\left\| hT(\lambda)K_2 - \mathrm{Op}_h(1-\chi^0) \right\|_{L^2(\Gamma)\to L^2(\Gamma)} \le C\delta \tag{7-8}$$

with a new constant $C > 0$ independent of λ and δ, where

$$K_2 := \mathrm{Op}_h(1-\chi^0)\mathrm{Op}_h(\tilde{\tau}_\delta)^{-1} = \mathcal{O}(1) : L^2(\Gamma) \to L^2(\Gamma).$$

By (7-3) and (7-8),

$$\left\| hT(\lambda)(K_1+K_2) - 1 \right\|_{L^2(\Gamma)\to L^2(\Gamma)} \le (C+1)\delta. \tag{7-9}$$

It follows from (7-9) that if δ is taken small enough, the operator $hT(\lambda)$ is invertible with an inverse satisfying the bound

$$\|(hT(\lambda))^{-1}\|_{L^2(\Gamma)\to L^2(\Gamma)} \le 2\|K_1\|_{L^2(\Gamma)\to L^2(\Gamma)} + 2\|K_2\|_{L^2(\Gamma)\to L^2(\Gamma)} \lesssim |\theta|^{-1} + 1. \tag{7-10}$$

It is easy to see that (7-10) implies (7-1). $\square$

References

[Cardoso et al. 2001] F. Cardoso, G. Popov, and G. Vodev, "Asymptotics of the number of resonances in the transmission problem", *Comm. Partial Differential Equations* **26**:9-10 (2001), 1811–1859. MR Zbl

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, Cambridge, 1999. MR Zbl

[Faierman 2014] M. Faierman, "The interior transmission problem: spectral theory", *SIAM J. Math. Anal.* **46**:1 (2014), 803–819. MR Zbl

[Galkowski 2015] J. Galkowski, "The quantum Sabine law for resonances in transmission problems", preprint, 2015. arXiv

[Hitrik et al. 2011] M. Hitrik, K. Krupchyk, P. Ola, and L. Päivärinta, "The interior transmission problem and bounds on transmission eigenvalues", *Math. Res. Lett.* **18**:2 (2011), 279–293. MR Zbl

[Lakshtanov and Vainberg 2013] E. Lakshtanov and B. Vainberg, "Applications of elliptic operator theory to the isotropic interior transmission eigenvalue problem", *Inverse Problems* **29**:10 (2013), art. id. 104003. MR Zbl

[Leung and Colton 2012] Y.-J. Leung and D. Colton, "Complex transmission eigenvalues for spherically stratified media", *Inverse Problems* **28**:7 (2012), art. id. 075005. MR Zbl

[Petkov 2016] V. Petkov, "Location of eigenvalues for the wave equation with dissipative boundary conditions", *Inverse Probl. Imaging* **10**:4 (2016), 1111–1139. MR Zbl

[Petkov and Vodev 2017a] V. Petkov and G. Vodev, "Asymptotics of the number of the interior transmission eigenvalues", *J. Spectr. Theory* **7**:1 (2017), 1–31. MR Zbl

[Petkov and Vodev 2017b] V. Petkov and G. Vodev, "Localization of the interior transmission eigenvalues for a ball", *Inverse Probl. Imaging* **11**:2 (2017), 355–372. MR Zbl

[Pham and Stefanov 2014] H. Pham and P. Stefanov, "Weyl asymptotics of the transmission eigenvalues for a constant index of refraction", *Inverse Probl. Imaging* **8**:3 (2014), 795–810. MR Zbl

[Robbiano 2013] L. Robbiano, "Spectral analysis of the interior transmission eigenvalue problem", *Inverse Problems* **29**:10 (2013), art. id. 104001. MR Zbl

[Robbiano 2016] L. Robbiano, "Counting function for interior transmission eigenvalues", *Math. Control Relat. Fields* **6**:1 (2016), 167–183. MR Zbl

[Sylvester 2013] J. Sylvester, "Transmission eigenvalues in one dimension", *Inverse Problems* **29**:10 (2013), art. id. 104009. MR Zbl

[Vodev 2015] G. Vodev, "Transmission eigenvalue-free regions", *Comm. Math. Phys.* **336**:3 (2015), 1141–1166. MR Zbl

[Vodev 2016] G. Vodev, "Transmission eigenvalues for strictly concave domains", *Math. Ann.* **366**:1-2 (2016), 301–336. MR Zbl

Received 17 Jan 2017. Revised 21 Jun 2017. Accepted 10 Aug 2017.

GEORGI VODEV: georgi.vodev@univ-nantes.fr
Université de Nantes, Laboratoire de Mathématiques Jean Leray, Nantes, France

ANALYSIS AND PDE
Vol. 11, No. 1, 2018
dx.doi.org/10.2140/apde.2018.11.237

HARDY–LITTLEWOOD INEQUALITIES ON COMPACT QUANTUM GROUPS OF KAC TYPE

SANG-GYUN YOUN

The Hardy–Littlewood inequality on the circle group $\mathbb{T}$ compares the L^p-norm of a function with a weighted ℓ^p-norm of its sequence of Fourier coefficients. The approach has recently been explored for compact homogeneous spaces and we study a natural analogue in the framework of compact quantum groups. In particular, in the case of the reduced group C^*-algebras and free quantum groups, we establish explicit $L^p - \ell^p$ inequalities through inherent information of the underlying quantum groups such as growth rates and the rapid decay property. Moreover, we show sharpness of the inequalities in a large class, including G a compact Lie group, $C_r^*(G)$ with G a polynomially growing discrete group and free quantum groups O_N^+, S_N^+.

1. Introduction

Hardy and Littlewood [1927] showed that there exists a constant C_p for each $1 < p \leq 2$ such that

$$\left(\sum_{n\in\mathbb{Z}} \frac{1}{(1+|n|)^{2-p}} |\hat{f}(n)|^p\right)^{\frac{1}{p}} \leq C_p \|f\|_{L^p(\mathbb{T})} \tag{1-1}$$

for all $f \in L^p(\mathbb{T})$, where $(\hat{f}(n))_{n\in\mathbb{Z}}$ is the sequence of Fourier coefficients of f.

This implies the multiplier

$$\mathcal{F}_w : L^p(\mathbb{T}) \to \ell^p(\mathbb{Z}), \quad f \mapsto (w(n)\hat{f}(n))_{n\in\mathbb{Z}}, \quad \text{with } w(n) := \frac{1}{(1+|n|)^{\frac{2-p}{p}}},$$

is bounded. Moreover, this is a stronger form of the Hardy–Littlewood–Sobolev embedding theorem

$$H_p^{\frac{1}{p}-\frac{1}{q}}(\mathbb{T}) \subseteq L^q(\mathbb{T}) \quad \text{for all } 1 < p < q < \infty,$$

where $H_p^s(\mathbb{T}) := \{f \in L^p(\mathbb{T}) : (1-\Delta)^{\frac{s}{2}}(f) \in L^p(\mathbb{T})\}$ is the Bessel potential space [Bényi and Oh 2013].

The Hardy–Littlewood inequality (1-1) has been studied on compact abelian groups by Hewitt and Ross [1974] and was recently extended to compact homogeneous manifolds by Akylzhanov, Nursultanov and Ruzhansky [Akylzhanov et al. 2015; 2016]. For compact Lie groups G with real dimension n, the

The author is supported by the TJ Park Science Fellowship and the Basic Science Research Program through the National Research Foundation of Korea (NRF), grant NRF-2015R1A2A2A01006882.
MSC2010: 20G42, 43A15, 46L51, 46L52.
Keywords: Hardy–Littlewood inequality, quantum groups, Fourier analysis.

2015 paper's result can be rephrased thus: for each $1 < p \leq 2$, there exists a constant $C_p > 0$ such that

$$\Big(\sum_{\pi \in \mathrm{Irr}(G)} \frac{1}{(1+\kappa_\pi)^{\frac{n(2-p)}{2}}} n_\pi^{2-\frac{p}{2}} \|\hat{f}(\pi)\|_{\mathrm{HS}}^p \Big)^{\frac{1}{p}} \leq C_p \|f\|_{L^p(G)} \quad \text{for all } f \in L^p(G). \tag{1-2}$$

Here, $\mathrm{Irr}(G)$ denotes a maximal family of mutually inequivalent irreducible unitary representations of G, $\|A\|_{\mathrm{HS}} := \mathrm{tr}(A^*A)^{\frac{1}{2}}$ and the Laplacian operator Δ on G satisfies $\Delta : \pi_{i,j} \mapsto -\kappa_\pi \pi_{i,j}$ for all $\pi = (\pi_{i,j})_{1 \leq i,j \leq n_\pi} \in \mathrm{Irr}(G)$ and all $1 \leq i, j \leq n_\pi$.

The left-hand side of the inequality (1-2) can be shown to dominate a more familiar quantity, which is a natural weighted ℓ^p-norm of its sequence of Fourier coefficients:

$$\Big(\sum_{\pi \in \mathrm{Irr}(G)} \frac{1}{(1+\kappa_\pi)^{\frac{n(2-p)}{2}}} n_\pi \|\hat{f}(\pi)\|_{S^p_{n_\pi}}^p \Big)^{\frac{1}{p}} \leq C_p \|f\|_{L^p(G)}. \tag{1-3}$$

Here, $\|A\|_{S^p_n} = \mathrm{tr}(|A|^p)^{\frac{1}{p}}$ is called the Schatten p-norm with respect to the (unnormalized) trace.

A notable point is that the Hardy–Littlewood inequalities on compact Lie groups (1-2) are determined by inherent geometric information, namely the real dimension and the natural length function on $\mathrm{Irr}(G)$. Indeed, $\pi \mapsto \sqrt{\kappa_\pi}$ is equivalent to the natural length $\|\cdot\|$ on $\mathrm{Irr}(G)$ (see Remark 6.1).

The main purpose of this paper is to establish new Hardy–Littlewood inequalities on compact quantum groups of Kac type by utilizing geometric information of the underlying quantum groups. As part of such efforts, we will present some explicit inequalities on important examples and such examples are listed as follows. The reduced group C^*-algebras $C_r^*(G)$ of discrete groups G, the free orthogonal quantum groups O_N^+ and the free permutation quantum groups S_N^+ are main targets. Of course, noncommutative L^p analysis on quantum groups is widely discussed from various perspectives [Caspers 2013; Franz et al. 2017; Junge et al. 2014; 2017; Wang 2017]. For the details of an operator algebraic approach to quantum groups themselves, see [Kustermans and Vaes 2000; 2003; Timmermann 2008; Woronowicz 1987].

In order to clarify our intention, let us show the main results of this paper on compact matrix quantum groups, which can be known to admit the natural length function $|\cdot| : \mathrm{Irr}(\mathbb{G}) \to \{0\} \cup \mathbb{N}$ (see Definition 3.3 and Proposition 3.4). The following inequalities are determined by inherent information of the underlying quantum groups, namely growth rates and the rapid decay property.

Theorem 1.1. *Let $\mathbb{G}$ be a compact matrix quantum group of Kac type and denote by $|\cdot|$ the natural length function on* $\mathrm{Irr}(\mathbb{G})$.

(1) *Let $\mathbb{G}$ have a polynomial growth with $\sum_{\alpha \in \mathrm{Irr}(\mathbb{G}) : |\alpha| \leq k} n_\alpha^2 \leq (1+k)^\gamma$ and $\gamma > 0$. Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(p)$ such that*

$$\Big(\sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} \frac{1}{(1+|\alpha|)^{(2-p)\gamma}} n_\alpha \|\hat{f}(\alpha)\|_{S^p_{n_\alpha}}^p \Big)^{\frac{1}{p}} \leq \Big(\sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} \frac{1}{(1+|\alpha|)^{(2-p)\gamma}} n_\alpha^{2-\frac{p}{2}} \|\hat{f}(\alpha)\|_{\mathrm{HS}}^p \Big)^{\frac{1}{p}} \leq K \|f\|_{L^p(\mathbb{G})} \tag{1-4}$$

for all $f \sim \sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} n_\alpha \,\mathrm{tr}(\hat{f}(\alpha) u^\alpha) \in L^p(\mathbb{G})$.

(2) *Let $\widehat{\mathbb{G}}$ have the rapid decay property with universal constants $C, \beta > 0$ such that*

$$\|f\|_{L^\infty(\mathbb{G})} \leq C(1+k)^\beta \|f\|_{L^2(\mathbb{G})}$$

for all $f \in \mathrm{span}(\{u^{\alpha}_{i,j} : |\alpha| = k,\ 1 \le i, j \le n_\alpha\})$. *Define*

$$s_k := \sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} n_\alpha^2.$$

Then, for each $1 < p \le 2$, *there exists a universal constant* $K = K(p)$ *such that*

$$\Bigg(\sum_{k\ge 0}\sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} \frac{1}{s_k^{\frac{(2-p)}{2}}(1+k)^{(2-p)(\beta+1)}} n_\alpha \|\hat{f}(\alpha)\|^p_{S^p_{n_\alpha}}\Bigg)^{\frac{1}{p}} \le \Bigg(\sum_{k\ge 0}\frac{1}{(1+k)^{(2-p)(\beta+1)}}\Bigg(\sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} n_\alpha\|\hat{f}(\alpha)\|^2_{\mathrm{HS}}\Bigg)^{\frac{p}{2}}\Bigg)^{\frac{1}{p}} \le K\|f\|_{L^p(\mathbb{G})} \quad (1\text{-}5)$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \,\mathrm{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

In particular, it is known that the rapid decay property of $\mathbb{F}_N$ can be strengthened in a general holomorphic setting [Kemp and Speicher 2007]. The improved result is called the strong Haagerup inequality. Based on these data, it can be shown that we can improve the Hardy–Littlewood inequality on $C^*_r(\mathbb{F}_N)$ by focusing on holomorphic forms. Theorem 5.3 justifies the claim and it seems appropriate to call the improved result a "strong Hardy–Littlewood inequality".

A natural perspective on the Hardy–Littlewood inequalities on compact Lie groups is that a properly chosen weight function $w : \mathrm{Irr}(G) \to (0,\infty)$ makes the corresponding multiplier

$$\mathcal{F}_w : L^p(G) \to \ell^p(\widehat{G}), \quad \text{given by } f \mapsto (w(\pi)\hat{f}(\pi))_{\pi\in\mathrm{Irr}(G)},$$

bounded for each $1 < p \le 2$. Indeed, our newly derived Hardy–Littlewood inequalities on compact quantum groups will give a specific pair (r, s) whose corresponding multiplier $\mathcal{F}_{w_{r,s}}$ is bounded, where

$$w_{r,s}(\alpha) := \frac{1}{r^{|\alpha|}(1+|\alpha|)^s}.$$

Moreover, in Section 6, we will show that there is no better pair (r', s') in that $\mathcal{F}_{r',s'}$ is unbounded whenever (1) $r' < r$ or (2) $r' = r,\ s' < s$ if $\mathbb{G}$ is one of the following: G a compact Lie group, $C^*_r(G)$ with polynomially growing discrete groups or one of the free quantum groups O_N^+, S_N^+. See Theorem 6.6.

This approach is quite natural because it is strongly related to Sobolev embedding properties. We will explore how they are related in Sections 6 and 7B. Indeed, for $G = \mathbb{T}^d$, we have $\mathcal{F}_{w_{0,s}} : L^p(\mathbb{T}^d) \to \ell^p(\mathbb{Z}^d)$ is bounded if and only if

$$H_q^{\frac{ps}{2-p}\left(\frac{1}{q}-\frac{1}{r}\right)}(\mathbb{T}^d) \subseteq L^r(\mathbb{T}^d) \quad \text{for all } 1 < q < r < \infty,$$

where $H^s_p(\mathbb{T}^d)$ is the Bessel potential space.

Lastly, in Section 7, we present some remarks that follow from this approach. We show that many free quantum groups do not admit infinite (central) Sidon sets and give a Sobolev embedding theorem-type interpretation of our results to $C(G)$ with compact Lie groups and $C^*_r(G)$ with polynomially growing discrete groups G. Also, we present an explicit inequality on quantum torus $\mathbb{T}^d_\theta$, which is not a quantum group [Sołtan 2010].

2. Preliminaries

2A. ***Compact quantum groups.*** A compact quantum group $\mathbb{G}$ is given by a unital C^*-algebra A and a unital $*$-homomorphism $\Delta : A \to A \otimes_{\min} A$ satisfying

(1) $(\Delta \otimes \mathrm{id}) \circ \Delta = (\mathrm{id} \otimes \Delta) \circ \Delta$;

(2) $\mathrm{span}\{\Delta(a)(b \otimes 1_A) : a, b \in A\}$ and $\mathrm{span}\{\Delta(a)(1_A \otimes b) : a, b \in A\}$ are dense in A.

Every compact quantum group admits a unique Haar state h on A such that

$$(h \otimes \mathrm{id})(\Delta(x)) = h(x)1_A = (\mathrm{id} \otimes h)(\Delta(x)) \quad \text{for all } x \in A.$$

A finite-dimensional representation of $\mathbb{G}$ is given by an element $u = (u_{i,j})_{1 \le i,j \le n} \in M_n(A)$ such that $\Delta(u_{i,j}) = \sum_{k=1}^n u_{i,k} \otimes u_{k,j}$ for all $1 \le i, j \le n$. We say that the representation u is unitary if $u^*u = uu^* = \mathrm{Id}_n \otimes 1_A \in M_n(A)$ and irreducible if $\{X \in M_n : Xu = uX\} = \mathbb{C} \cdot \mathrm{Id}_n$, where Id_n is the identity matrix in M_n.

We say that finite-dimensional unitary representations u_1 and u_2 are equivalent if there exists a unitary matrix X such that $u_1 X = X u_2$ and let $\{u^\alpha = (u^\alpha_{i,j})_{1 \le i,j \le n_\alpha}\}_{\alpha \in \mathrm{Irr}(\mathbb{G})}$ be a maximal family of mutually inequivalent finite-dimensional unitary irreducible representations of $\mathbb{G}$. It is well known that there is a unique positive invertible matrix $Q_\alpha \in M_{n_\alpha}$ for each $\alpha \in \mathrm{Irr}(\mathbb{G})$ such that $\mathrm{tr}(Q_\alpha) = \mathrm{tr}(Q_\alpha^{-1})$ and

$$h((u^\beta_{s,t})^* u^\alpha_{i,j}) = \frac{\delta_{\alpha,\beta}\delta_{j,t}(Q_\alpha^{-1})_{i,s}}{\mathrm{tr}(Q_\alpha)} \quad \text{for all } \alpha, \beta \in \mathrm{Irr}(\mathbb{G}),\ 1 \le i, j \le n_\alpha,\ 1 \le s, t \le n_\beta,$$
$$h(u^\beta_{s,t}(u^\alpha_{i,j})^*) = \frac{\delta_{\alpha,\beta}\delta_{i,s}(Q_\alpha)_{j,t}}{\mathrm{tr}(Q_\alpha)} \quad \text{for all } \alpha, \beta \in \mathrm{Irr}(\mathbb{G}),\ 1 \le i, j \le n_\alpha,\ 1 \le s, t \le n_\beta.$$

We say that $\mathbb{G}$ is of Kac type if $Q_\alpha = \mathrm{Id}_{n_\alpha} \in M_{n_\alpha}$ for all $\alpha \in \mathrm{Irr}(\mathbb{G})$. In this case, the Haar state h is tracial.

Lastly, we define $C_r(\mathbb{G})$ as the image of A in the GNS representation with respect to the Haar state h and $L^\infty(\mathbb{G}) := C_r(\mathbb{G})''$. The Haar state h has a normal faithful extension to $L^\infty(\mathbb{G})$.

2B. ***Noncommutative L^p-spaces.*** Let $\mathcal{M}$ be a von Neumann algebra with a normal faithful tracial state ϕ. Note that the von Neumann algebra $\mathcal{M}$ admits the unique predual $\mathcal{M}_*$. We define $L^1(\mathcal{M}, \phi) := \mathcal{M}_*$ and $L^\infty(\mathcal{M}, \phi) := \mathcal{M}$, and then consider a contractive injection $j : \mathcal{M} \to \mathcal{M}_*$, given by $[j(x)](y) := h(yx)$ for all $y \in \mathcal{M}$. The map j has dense range.

Now $(\mathcal{M}, \mathcal{M}_*)$ is a compatible pair of Banach spaces and for all $1 < p < \infty$, we can define noncommutative L^p-space $L^p(\mathcal{M}, \phi) := (\mathcal{M}, \mathcal{M}_*)_{\frac{1}{p}}$, where $(\cdot, \cdot)_{\frac{1}{p}}$ is the complex interpolation space. For any $x \in L^\infty(\mathcal{M}, \phi)$, its L^p-norm, for all $1 \le p < \infty$, is given by

$$\|x\|_{L^p(\mathcal{M},\phi)} = \phi(|x|^p)^{\frac{1}{p}}.$$

In particular, for all $1 \le p \le \infty$, we denote by $L^p(\mathbb{G})$ the noncommutative L^p-space associated with the von Neumann algebra $L^\infty(\mathbb{G})$ of a compact quantum group $\mathbb{G}$ of Kac type and the tracial Haar state h. Then the space of polynomials

$$\mathrm{Pol}(\mathbb{G}) := \mathrm{span}\big(\{u^\alpha_{i,j} : \alpha \in \mathrm{Irr}(\mathbb{G}),\ 1 \le i, j \le n_\alpha\}\big)$$

is dense in $C_r(\mathbb{G})$ and $L^p(\mathbb{G})$ for all $1 \le p < \infty$.

Under the assumption that $\mathbb{G}$ is of Kac type, for $1 \le p < \infty$,

$$\ell^p(\widehat{\mathbb{G}}) := \left\{ (A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})} \in \prod_{\alpha \in \mathrm{Irr}(\mathbb{G})} M_{n_\alpha} : \sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(|A_\alpha|^p) < \infty \right\}$$

and the natural ℓ^p-norm of $(A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})} \in \ell^p(\widehat{\mathbb{G}})$ is defined by

$$\|(A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})}\|_{\ell^p(\widehat{\mathbb{G}})} := \left(\sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(|A_\alpha|^p) \right)^{\frac{1}{p}} = \left(\sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} n_\alpha \|A_\alpha\|_{S^p_{n_\alpha}}^p \right)^{\frac{1}{p}}.$$

Also,

$$\ell^\infty(\widehat{\mathbb{G}}) := \left\{ (A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})} \in \prod_{\alpha \in \mathrm{Irr}(\mathbb{G})} M_{n_\alpha} : \sup_{\alpha \in \mathrm{Irr}(\mathbb{G})} \|A_\alpha\| < \infty \right\}$$

and the ℓ^∞-norm of $(A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})} \in \ell^\infty(\widehat{\mathbb{G}})$ is defined by

$$\|(A_\alpha)_{\alpha \in \mathrm{Irr}(\mathbb{G})}\|_{\ell^\infty(\widehat{\mathbb{G}})} := \sup_{\alpha \in \mathrm{Irr}(\mathbb{G})} \|A_\alpha\|.$$

It is known that $\ell^1(\widehat{\mathbb{G}}) = (\ell^\infty(\widehat{\mathbb{G}}))_*$ and $\ell^p(\widehat{\mathbb{G}}) = \big(\ell^\infty(\widehat{\mathbb{G}}), \ell^1(\widehat{\mathbb{G}})\big)_{\frac{1}{p}}$ for all $1 < p < \infty$. For details and more general framework of noncommutative L^p theory, see [Haagerup 1979; Pisier and Xu 2003; Xu 2007].

2C. *Fourier analysis on compact quantum groups.* For a compact quantum group $\mathbb{G}$, the Fourier transform $\mathcal{F} : L^1(\mathbb{G}) \to \ell^\infty(\widehat{\mathbb{G}})$, $\psi \mapsto \widehat{\psi}$, is defined by

$$(\widehat{\psi}(\alpha))_{i,j} := \psi((u^\alpha_{j,i})^*) \quad \text{for all } \alpha \in \mathrm{Irr}(\mathbb{G}),\ 1 \le i, j \le n_\alpha.$$

It is also known that $\mathcal{F}$ is an injective contractive map and it is an isometry from $L^2(\mathbb{G})$ onto $\ell^2(\widehat{\mathbb{G}})$ [Wang 2017, Propositions 3.1 and 3.2]. Then, by the interpolation theorem, we are able to induce the Hausdorff–Young inequality again; i.e., $\mathcal{F}$ is a contractive map from $L^p(\mathbb{G})$ into $\ell^{p'}(\widehat{\mathbb{G}})$ for each $1 \le p \le 2$, where p' is the conjugate of p.

We define the Fourier series of $f \in L^1(\mathbb{G})$ by $\sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} d_\alpha \operatorname{tr}(\hat{f}(\alpha) Q_\alpha u^\alpha)$ and denote it by $f \sim \sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} d_\alpha \operatorname{tr}(\hat{f}(\alpha) Q_\alpha u^\alpha)$. In particular, if $\mathbb{G}$ is of Kac type, the Fourier series is of the form $f \sim \sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha) u^\alpha)$.

2D. *The reduced group C^*-algebras.* The reduced group C^*-algebra $C^*_r(G)$ can be defined for all locally compact groups, but we only consider discrete groups in this paper since we aim to regard it as a compact quantum group.

Definition 2.1. Let G be a discrete group and define $\lambda_g \in B(\ell^2(G))$ for each $g \in G$ by

$$[(\lambda_g)(f)](x) := f(g^{-1}x) \quad \text{for all } x \in G.$$

Then the reduced group C^*-algebra $C^*_r(G)$ is defined as the norm-closure of the space $\mathrm{span}(\{\lambda_g : g \in G\})$ in $B(\ell^2(G))$. Moreover, we define a comultiplication $\Delta : C^*_r(G) \to C^*_r(G) \otimes_{\min} C^*_r(G)$ by $\lambda_g \mapsto \lambda_g \otimes \lambda_g$ for all $g \in G$. Then $(C^*_r(G), \Delta)$ forms a compact quantum group.

Note that, for $\mathbb{G} = (C^*_r(G), \Delta)$ of a discrete group G, it is of Kac type and $L^\infty(\mathbb{G})$ is nothing but the group von Neumann algebra $\mathrm{VN}(G)$ and $\mathrm{Irr}(\mathbb{G}) = \{\lambda_g\}_{g \in G}$ can be identified with G. In this case, we use the notation $L^p(\mathrm{VN}(G)) = L^p(\mathbb{G})$ conventionally. In particular, $L^1(\mathrm{VN}(G))$ is called the Fourier algebra, denoted by $A(G)$. It is well known that $A(G)$ embeds contractively into $C_0(G)$, so that $A(G)$ can be considered as a function space on G.

2E. ***Free quantum groups of Kac type.***

Definition 2.2 (free orthogonal quantum group [Wang 1995]). Let $N \geq 2$ and A be the universal unital C^*-algebra, which is generated by the N^2 self-adjoint elements $u_{i,j}$ with $1 \leq i, j \leq N$ satisfying the relations

$$\sum_{k=1}^{N} u_{k,i}u_{k,j} = \sum_{k=1}^{N} u_{i,k}u_{j,k} = \delta_{i,j} \quad \text{for all } 1 \leq i, j \leq N.$$

Also, we define a comultiplication $\Delta : A \to A \otimes_{\min} A$ by $u_{i,j} \mapsto \sum_{k=1}^{N} u_{i,k} \otimes u_{k,j}$. Then (A, Δ) forms a compact quantum group called the free orthogonal quantum group. We denote it by O_N^+.

Definition 2.3 (free permutation quantum group [Wang 1998]). Let $N \geq 2$ and A be the universal unital C^*-algebra generated by the N^2 self-adjoint elements $u_{i,j}$ with $1 \leq i, j \leq N$ satisfying the relations

$$u_{i,j}^2 = u_{i,j} = u_{i,j}^* \quad \text{and} \quad \sum_{k=1}^{N} u_{i,k} = \sum_{k=1}^{N} u_{k,j} = 1_A \quad \text{for all } 1 \leq i, j \leq N.$$

Also, we define a comultiplication $\Delta : A \to A \otimes_{\min} A$ by $u_{i,j} \mapsto \sum_{k=1}^{N} u_{i,k} \otimes u_{k,j}$. Then (A, Δ) forms a compact quantum group called the free permutation quantum group. We denote it by S_N^+.

These free quantum groups are of Kac type, so that the Haar states are tracial states. Also, for all $N \geq 2$, the families $\mathrm{Irr}(O_N^+)$ and $\mathrm{Irr}(S_{N+2}^+)$ can be identified with $\{0\} \cup \mathbb{N}$. Moreover,

$$\begin{aligned} n_k &= k+1 && \text{if } \mathbb{G} = O_2^+, \\ n_k &= 2k+1 && \text{if } \mathbb{G} = S_4^+, \\ n_k &\approx r_0^k && \text{if } \mathbb{G} = O_N^+ \text{ or } S_{N+2}^+ \text{ with } N \geq 3, \end{aligned}$$

where r_0 is the largest solution of the equation $X^2 - NX + 1 = 0$ [Banica and Vergnioux 2009].

2F. ***The noncommutative Marcinkiewicz interpolation theorem.*** The classical Marcinkiewicz interpolation theorem [Folland 1999, Theorem 6.28] has a natural noncommutative analogue for semifinite von Neumann algebras. Throughout this paper, we say that a map $T : L^p(M) \to L^q(N)$ is sublinear if for any $x, y \in L^p(M)$ and $\alpha \in \mathbb{C}$

$$|T(x+y)| \leq |T(x)| + |T(y)| \quad \text{and} \quad |T(\alpha x)| = |\alpha||T(x)|.$$

The following theorem is a special case of [Bekjan and Chen 2012, Theorem 2.1]. Denote by $L^0(M)$ the topological $*$-algebra of measurable operators with respect to (M, ϕ).

Theorem 2.4 (the noncommutative Marcinkiewicz interpolation theorem). *Let M and N be von Neumann algebras equipped with normal semifinite faithful traces ϕ and ψ respectively and let* $1 \leq p_1 < p < p_2 < \infty$.

Assume that a sublinear map $A : L^0(M) \to L^0(N)$ satisfies the following conditions: there exist $C_1, C_2 > 0$ such that for any $T_1 \in L^{p_1}(M)$, $T_2 \in L^{p_2}(M)$ and any $y > 0$,

$$\psi\big(1_{(y,\infty)}(|AT_1|)\big) \le \left(\frac{C_1}{y}\right)^{p_1} \|T_1\|_{L^{p_1}(M)}^{p_1}, \quad \psi\big(1_{(y,\infty)}(|AT_2|)\big) \le \left(\frac{C_2}{y}\right)^{p_2} \|T_2\|_{L^{p_2}(M)}^{p_2}. \tag{2-1}$$

Then $A : L^p(M) \to L^p(N)$ is a bounded map.

Proof. Choose a specific Orlicz function $\Phi(t) = t^p$. Then $p_\Phi = q_\Phi = p$ under the notation of [Bekjan and Chen 2012, Theorem 2.1]. □

If the sublinear operator A satisfies the left condition of the inequality (2-1), then we say that A is of weak type (p_1, p_1). Also, the boundedness of $A : L^p(M) \to L^p(N)$ implies A is of weak type (p, p).

Now denote the space of all functions on the discrete space $\mathrm{Irr}(\mathbb{G})$ by $c(\mathrm{Irr}(\mathbb{G}), \nu)$ with a positive measure ν. Then the theorem above is reformulated as follows:

Corollary 2.5. *Let $\mathbb{G}$ be a compact quantum group of Kac type and let $1 \le p_1 < p < p_2 < \infty$. Assume that $A : L^\infty(\mathbb{G}) \to c(\mathrm{Irr}(\mathbb{G}), \nu)$ is sublinear and satisfies the following conditions: there exist $C_1, C_2 > 0$ such that for any $T_1 \in L^{p_1}(\mathbb{G})$, $T_2 \in L^{p_2}(\mathbb{G})$ and any $y > 0$,*

$$\sum_{\alpha:\ |(AT_1)(\alpha)|\ge y} \nu(\alpha) \le \left(\frac{C_1}{y}\right)^{p_1} \|T_1\|_{L^{p_1}(\mathbb{G})}^{p_1}, \qquad \sum_{\alpha:\ |(AT_2)(\alpha)|\ge y} \nu(\alpha) \le \left(\frac{C_2}{y}\right)^{p_2} \|T_2\|_{L^{p_2}(\mathbb{G})}^{p_2}.$$

Then $A : L^p(\mathbb{G}) \to \ell^p(\mathrm{Irr}(\mathbb{G}), \nu)$ is a bounded map.

3. Paley-type inequalities

3A. ***General approach.*** In this subsection, a Paley-type inequality is derived for compact quantum groups of Kac type by employing fundamental techniques such as the Hausdorff–Young inequality, the Plancherel theorem and the noncommutative Marcinkiewicz interpolation theorem.

We prove the following theorem by adapting techniques used in [Akylzhanov et al. 2015].

Theorem 3.1. *Let $\mathbb{G}$ be a compact quantum group of Kac type and let $w : \mathrm{Irr}(\mathbb{G}) \to (0, \infty)$ be a function such that $C_w := \sup_{t>0}\{t \cdot \sum_{\alpha:\ w(\alpha)\ge t} n_\alpha^2\} < \infty$. Then, for each $1 < p \le 2$, there exists a universal constant $K = K(p) > 0$ such that*

$$\left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} w(\alpha)^{2-p} n_\alpha^{2-\frac{p}{2}} \|\hat{f}(\alpha)\|_{\mathrm{HS}}^p\right)^{\frac{1}{p}} \le K\|f\|_{L^p(\mathbb{G})} \tag{3-1}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

Proof. Put $\nu(\alpha) := w(\alpha)^2 n_\alpha^2$. We will show that the sublinear operator

$$A : L^1(\mathbb{G}) \to c(\mathrm{Irr}(\mathbb{G}), \nu), \quad f \mapsto \left(\frac{\|\hat{f}(\alpha)\|_{\mathrm{HS}}}{\sqrt{n_\alpha}\, w(\alpha)}\right)_{\alpha\in\mathrm{Irr}(\mathbb{G})},$$

is a well-defined bounded map from $L^p(\mathbb{G})$ into $\ell^p(\mathrm{Irr}(\mathbb{G}), \nu)$ for all $1 < p \le 2$.

First of all,

$$\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} \|(Af)(\alpha)\|_{\mathrm{HS}}^2 \nu(\alpha) = \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \|\hat{f}(\alpha)\|_{\mathrm{HS}}^2 = \|f\|_{L^2(\mathbb{G})}^2.$$

This implies $A : L^2(\mathbb{G}) \to \ell^2(\mathrm{Irr}(\mathbb{G}), \nu)$ is an isometry.

Second, for all $y > 0$, since

$$\frac{\|\hat{f}(\alpha)\|_{\mathrm{HS}}}{\sqrt{n_\alpha}} = \left(\frac{\mathrm{tr}(\hat{f}(\alpha)^*\hat{f}(\alpha))}{n_\alpha}\right)^{\frac{1}{2}} \le \|\hat{f}(\alpha)\| \le \|f\|_{L^1(\mathbb{G})},$$

we have

$$\begin{aligned}
\sum_{\alpha:\ \|Af(\alpha)\|_{\mathrm{HS}}\ge y} \nu(\alpha) &\le \sum_{\alpha:\ w(\alpha)\le \frac{\|f\|_1}{y}} w(\alpha)^2 n_\alpha^2 = \sum_{\alpha:\ w(\alpha)\le\frac{\|f\|_1}{y}} \int_0^{w(\alpha)^2} n_\alpha^2\, dx \\
&= \int_0^{\left(\frac{\|f\|_1}{y}\right)^2} \Bigg(\sum_{\alpha:\ x^{\frac{1}{2}}\le w(\alpha)\le \frac{\|f\|_1}{y}} n_\alpha^2\Bigg)\, dx \quad \text{(by the Fubini theorem)} \\
&= 2\int_0^{\frac{\|f\|_1}{y}} t\Bigg(\sum_{\alpha:\ t\le w(\alpha)\le\frac{\|f\|_1}{y}} n_\alpha^2\Bigg)\, dt \quad \text{(by substituting } x \text{ to } t^2\text{)} \\
&\le 2C_w \frac{\|f\|_1}{y}.
\end{aligned}$$

This shows that A is of weak type $(1,1)$ with $C_1 = 2C_w$.

Now, by Corollary 2.5,

$$\left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} w(\alpha)^{2-p} n_\alpha^{p\left(\frac{2}{p}-\frac{1}{2}\right)} \|\hat{f}(\alpha)\|_{\mathrm{HS}}^p\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})}. \qquad \square$$

The left-hand side of the inequality (3-1) dominates a more familiar quantity, which is a natural weighted ℓ^p-norm of the sequence of Fourier coefficients $(\hat{f}(\alpha))_{\alpha\in\mathrm{Irr}(\mathbb{G})}$. Recall that the natural noncommutative ℓ^p-norm on $\ell^\infty(\widehat{\mathbb{G}}) = \ell^\infty - \bigoplus_{\alpha\in\mathrm{Irr}(\mathbb{G})} M_{n_\alpha}$ is given by

$$\|(A_\alpha)_{\alpha\in\mathrm{Irr}(\mathbb{G})}\|_{\ell^p(\widehat{\mathbb{G}})} = \left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \|A_\alpha\|_{S_{n_\alpha}^p}^p\right)^{\frac{1}{p}}$$

under the condition where $\mathbb{G}$ is of Kac type.

Corollary 3.2. *Let $1 < p \le 2$ and w be a function which satisfies the condition of Theorem 3.1. Then we have that*

$$\left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} w(\alpha)^{2-p} n_\alpha \|\hat{f}(\alpha)\|_{S_{n_\alpha}^p}^p\right)^{\frac{1}{p}} \le K\|f\|_{L^p(\mathbb{G})}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha\, \mathrm{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

Proof. First of all,

$$\mathrm{tr}(|\hat{f}(\alpha)|^p) = \|\hat{f}(\alpha)\|^p_{S^p_{n_\alpha}}.$$

Put $\frac{1}{r} = \frac{1}{p} - \frac{1}{2}$. Then $2 < r \le \infty$ and

$$\mathrm{tr}(|\hat{f}(\alpha)|^p) \le \|\hat{f}(\alpha)\|^p_{\mathrm{HS}} \, \|\mathrm{Id}_{n_\alpha}\|^p_{S^r_{n_\alpha}} = n_\alpha^{1-\frac{p}{2}} \|\hat{f}(\alpha)\|^p_{\mathrm{HS}}. \qquad \square$$

Now we discuss an important subclass of compact quantum groups, namely compact matrix quantum groups which admit the natural length function on $\mathrm{Irr}(\mathbb{G})$.

Definition 3.3. A compact matrix quantum group is given by a triple (A, Δ, u), where A is a unital C^*-algebra A, $\Delta : A \to A \otimes_{\min} A$ is a $*$-homomorphism and $u = (u_{i,j})_{1\le i,j\le n} \in M_n(A)$ is a unitary such that

(1) $\Delta : u_{i,j} \mapsto \sum_{k=1}^n u_{i,k} \otimes u_{k,j}$,

(2) $\bar{u} = (u^*_{i,j})_{1\le i,j\le n}$ is invertible in $M_n(A)$,

(3) $\{u_{i,j}\}_{1\le i,j\le n}$ generates A as a C^*-algebra.

By definition, the free orthogonal quantum groups O_N^+ and the free permutation quantum groups S_N^+ are compact matrix quantum groups. Also, in the class of compact quantum groups, the subclass of compact matrix quantum groups is characterized by the following proposition. The conjugate $\bar{\alpha} \in \mathrm{Irr}(\mathbb{G})$ of $\alpha \in \mathrm{Irr}(\mathbb{G})$ is determined by $u^{\bar{\alpha}} := Q_\alpha^{\frac{1}{2}} \bar{u}^\alpha Q_\alpha^{-\frac{1}{2}}$.

Proposition 3.4 [Timmermann 2008]. *A compact quantum group is a compact matrix quantum group if and only if there exists a finite set $S := \{\alpha_1, \dots, \alpha_n\} \subseteq \mathrm{Irr}(\mathbb{G})$ such that any $\alpha \in \mathrm{Irr}(\mathbb{G})$ is contained in some iterated tensor product of elements $\alpha_1, \bar{\alpha}_1, \dots, \alpha_n, \bar{\alpha}_n$ and the trivial representation.*

Then there is a natural way to define a length function on $\mathrm{Irr}(\mathbb{G})$ [Vergnioux 2007]. For nontrivial $\alpha \in \mathrm{Irr}(\mathbb{G})$, the natural length $|\alpha|$ is defined by

$$\min\bigl\{m \in \mathbb{N} : \exists \beta_1, \dots, \beta_m \text{ such that } \alpha \subseteq \beta_1 \otimes \cdots \otimes \beta_m,\ \beta_j \in \{\alpha_k, \bar{\alpha}_k\}_{k=1}^n\bigr\}.$$

The length of the trivial representation is defined by 0.

Then it becomes possible to extract explicit inequalities from Theorem 3.1 and Corollary 3.2 by inserting geometric information of the underlying quantum groups, namely growth rates that are estimated by the quantities $b_k := \sum_{|\alpha|\le k} n_\alpha^2$ [Banica and Vergnioux 2009].

Corollary 3.5. *Let a compact matrix quantum group $\mathbb{G}$ of Kac type satisfy*

$$b_k = \sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|\le k}} n_\alpha^2 \le C(1+k)^\gamma \quad \textit{for all } k \ge 0 \textit{ with } C, \gamma > 0$$

with respect to the natural length function. Then, for each $1 < p \le 2$, there exists a universal constant $K = K(p)$ such that

$$\left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} \frac{1}{(1+|\alpha|)^{(2-p)\gamma}} n_\alpha \|\hat{f}(\alpha)\|^p_{S^p_{n_\alpha}}\right)^{\frac{1}{p}} \le \left(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} \frac{1}{(1+|\alpha|)^{(2-p)\gamma}} n_\alpha^{2-\frac{p}{2}} \|\hat{f}(\alpha)\|^p_{\mathrm{HS}}\right)^{\frac{1}{p}} \le K\|f\|_{L^p(\mathbb{G})} \tag{3-2}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \,\mathrm{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

Proof. Consider the weight function $w(\alpha) := 1/(1+|\alpha|)^\gamma$. Then

$$\sup_{t>0}\Big\{t\cdot\sum_{\alpha:|\alpha|\le t^{-1/\gamma}-1} n_\alpha^2\Big\} = \sup_{0<t\le 1}\Big\{t\cdot\sum_{\alpha:|\alpha|\le t^{-1/\gamma}-1} n_\alpha^2\Big\} \le C \sup_{0<t\le 1} t\cdot (t^{-\frac{1}{\gamma}})^\gamma = C.$$

Now the conclusion is reached by Theorem 3.1 and Corollary 3.2. □

3B. ***A Paley-type inequality under the rapid decay property.*** In this subsection, we still assume that $\mathbb{G}$ is a compact matrix quantum group of Kac type. One of the major observations of this paper is that more detailed geometric information improves Theorem 3.1 and Corollary 3.2 in various "exponentially growing" cases. A more refined Paley-type inequality can be obtained under the condition that $\widehat{\mathbb{G}}$ has the rapid decay property in the sense of [Vergnioux 2007].

Definition 3.6 [Vergnioux 2007]. Let $\mathbb{G}$ be a compact matrix quantum group of Kac type. Then we say that $\widehat{\mathbb{G}}$ has the rapid decay property with respect to the natural length function on $\mathrm{Irr}(\mathbb{G})$ if there exist $C, \beta > 0$ such that

$$\Big\| \sum_{\substack{\alpha\in \mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} \sum_{i,j=1}^{n_\alpha} a_{i,j}^\alpha u_{i,j}^\alpha \Big\|_{L^\infty(\mathbb{G})} \le C(1+k)^\beta \Big\| \sum_{\substack{\alpha\in \mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} \sum_{i,j=1}^{n_\alpha} a_{i,j}^\alpha u_{i,j}^\alpha \Big\|_{L^2(\mathbb{G})} \tag{3-3}$$

for any $k \ge 0$ and scalars $a_{i,j}^\alpha \in \mathbb{C}$.

Notation. (1) When the natural length function on $\mathrm{Irr}(\mathbb{G})$ is given, we use the notation

$$S_k := \{\alpha \in \mathrm{Irr}(\mathbb{G}) : |\alpha| = k\} \quad \text{and} \quad s_k := \sum_{\alpha\in S_k} n_\alpha^2.$$

(2) We denote by p_k the orthogonal projection from $L^2(\mathbb{G})$ to the closure of

$$\mathrm{span}\big(\{u_{i,j}^\alpha : \alpha \in S_k,\ 1 \le i, j \le n_\alpha\}\big).$$

Proposition 3.7. *Suppose a compact matrix quantum group $\mathbb{G}$ is of Kac type and $\widehat{\mathbb{G}}$ has the rapid decay property with respect to the natural length function on* $\mathrm{Irr}(\mathbb{G})$ *and with inequality* (3-3)*. Then we have*

$$\sup_{k\ge 0} \frac{\big(\sum_{\alpha\in\mathrm{Irr}(\mathbb{G}):|\alpha|=k} n_\alpha \|\hat f(\alpha)\|_{\mathrm{HS}}^2\big)^{\frac12}}{(k+1)^\beta} \le C\|f\|_{L^1(\mathbb{G})} \quad \textit{for all } f \in L^1(\mathbb{G}). \tag{3-4}$$

Proof. Since $L^1(\mathbb{G})$ is isometrically embedded into the dual space $M(\mathbb{G}) := C_r(\mathbb{G})^*$ and $\mathrm{Pol}(\mathbb{G})$ is dense in $C_r(\mathbb{G})$, we have

$$\begin{aligned}\|f\|_{L^1(\mathbb{G})} &= \sup_{\substack{x\in \mathrm{Pol}(\mathbb{G})\\ \|x\|_{L^\infty(\mathbb{G})}\le 1}} \langle f, x\rangle_{L^1(\mathbb{G}),L^\infty(\mathbb{G})} = \sup_{\substack{x\in \mathrm{Pol}(\mathbb{G})\\ \|x\|_{L^\infty(\mathbb{G})}\le 1}} \langle f, x^*\rangle_{L^1(\mathbb{G}),L^\infty(\mathbb{G})}\\ &= \sup_{\substack{x\in \mathrm{Pol}(\mathbb{G})\\ \|x\|_{L^\infty(\mathbb{G})}\le 1}} \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \,\mathrm{tr}(\hat f(\alpha)\hat x(\alpha)^*)\end{aligned}$$

$$\geq \sup_{\substack{x\in \mathrm{Pol}(\mathbb{G}) \\ \sum_{k\geq 0} C(k+1)^{\beta}\left(\sum_{\alpha:|\alpha|=k} n_\alpha \|\hat{x}(\alpha)\|_{\mathrm{HS}}^2\right)^{1/2}\leq 1}} \sum_{\alpha\in \mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha)\hat{x}(\alpha)^*)$$

$$\geq \sup_{k\geq 0} \sup_{\left(\sum_{\alpha:|\alpha|=k} n_\alpha \|\hat{x}(\alpha)\|_{\mathrm{HS}}^2\right)^{1/2}\leq 1} \sum_{\substack{\alpha\in \mathrm{Irr}(\mathbb{G}) \\ |\alpha|=k}} \frac{n_\alpha}{C(k+1)^{\beta}} \operatorname{tr}(\hat{f}(\alpha)\hat{x}(\alpha)^*)$$

$$= \sup_{k\geq 0} \frac{\left(\sum_{\alpha\in \mathrm{Irr}(\mathbb{G}):|\alpha|=k} n_\alpha \|\hat{f}(\alpha)\|_{\mathrm{HS}}^2\right)^{\frac{1}{2}}}{C(k+1)^{\beta}}. \tag{3-5}$$

This completes the proof. □

Theorem 3.8. *Let a compact matrix quantum group $\mathbb{G}$ be of Kac type and $\widehat{\mathbb{G}}$ have the rapid decay property with respect to the natural length function on* $\mathrm{Irr}(\mathbb{G})$ *and with inequality* (3-3). *Also, suppose that a weight function $w : \{0\}\cup\mathbb{N}\to(0,\infty)$ satisfies*

$$C_w := \sup_{y>0}\left\{ y\cdot \sum_{k\geq 0:\frac{(k+1)^{\beta}}{w(k)}\leq \frac{1}{y}} (k+1)^{2\beta}\right\} < \infty. \tag{3-6}$$

Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(p) > 0$ such that

$$\left(\sum_{k\geq 0} w(k)^{2-p}\left(\sum_{\substack{\alpha\in \mathrm{Irr}(\mathbb{G}) \\ |\alpha|=k}} n_\alpha \|\hat{f}(\alpha)\|_{\mathrm{HS}}^2\right)^{\frac{p}{2}}\right)^{\frac{1}{p}} \leq K\|f\|_{L^p(\mathbb{G})} \tag{3-7}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha)u^{\alpha}) \in L^p(\mathbb{G})$.

Proof. Put $\nu(k) := w(k)^2$. We will show that the sublinear operator

$$A : L^1(\mathbb{G}) \to c(\{0\}\cup\mathbb{N},\nu), \quad f \mapsto \left(\frac{\|p_k(f)\|_{L^2(\mathbb{G})}}{w(k)}\right)_{k\geq 0},$$

is a well-defined bounded map from $L^p(\mathbb{G})$ into $\ell^p(\{0\}\cup\mathbb{N},\nu)$ for all $1 < p \leq 2$.

First of all,

$$\sum_{k\geq 0}\left(\frac{\|p_k(f)\|_{L^2(\mathbb{G})}}{w(k)}\right)^2 \nu(k) = \sum_{k\geq 0} \|p_k(f)\|_{L^2(\mathbb{G})}^2 = \|f\|_{L^2(\mathbb{G})}^2.$$

Therefore, $A : L^2(\mathbb{G}) \to \ell^2(\{0\}\cup\mathbb{N},\nu)$ is an isometry.

Secondly, for all $y > 0$,

$$\sum_{\substack{k\geq 0 \\ (Af)(k)>y}} \nu(k) \leq \sum_{k:\frac{w(k)}{(k+1)^{\beta}}<\frac{C\|f\|_{L^1(\mathbb{G})}}{y}} w(k)^2$$

by Proposition 3.7.

Now put $\tilde{w}(k) := w(k)/(k+1)^{\beta}$. Then

$$\sum_{k:\tilde{w}(k)<\frac{C\|f\|_{L^1(\mathbb{G})}}{y}} \int_0^{\tilde{w}(k)^2} (k+1)^{2\beta}\,dx \leq \int_0^{\left(\frac{C\|f\|_{L^1(\mathbb{G})}}{y}\right)^2} \sum_{k:\sqrt{x}\leq\tilde{w}(k)} (k+1)^{2\beta}\,dx$$

$$= 2\int_0^{\frac{C\|f\|_{L^1(\mathbb{G})}}{y}} t\cdot \sum_{k:t\leq\tilde{w}(k)} (k+1)^{2\beta}\,dt \quad \text{(by substituting } x=t^2\text{)}$$

$$\leq \frac{2C_w C\|f\|_{L^1(\mathbb{G})}}{y}.$$

Therefore, by Corollary 2.5, we can obtain

$$\left(\sum_{k\geq 0} w(k)^{2-p}\|p_k(f)\|_{L^2(\mathbb{G})}^p\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})}. \qquad \square$$

Corollary 3.9. *Let a compact matrix quantum group $\mathbb{G}$ be of Kac type and $\widehat{\mathbb{G}}$ have the rapid decay property with respect to the natural length function on* Irr($\mathbb{G}$) *and with inequality* (3-3)*. Then, for each* $1 < p \leq 2$, *there exists a universal constant* $K = K(p) > 0$ *such that*

$$\left(\sum_{k\geq 0} \frac{1}{(1+k)^{(2-p)(\beta+1)}} \left(\sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} n_\alpha \|\hat{f}(\alpha)\|_{\mathrm{HS}}^2\right)^{\frac{p}{2}}\right)^{\frac{1}{p}} \leq K\|f\|_{L^p(\mathbb{G})} \tag{3-8}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

Proof. Take $w(k) := 1/(1+k)^{\beta+1}$. Then

$$C_w = \sup_{y>0}\left\{ y\cdot \sum_{k\geq 0:(1+k)^{2\beta+1}\leq\frac{1}{y}} (1+k)^{2\beta}\right\}$$

$$\leq \sup_{0<y\leq 1}\left\{ y\cdot \int_1^{\left(\frac{1}{y}\right)^{1/(2\beta+1)}+1} t^{2\beta}\,dt\right\} \leq \sup_{0<y\leq 1}\left\{ y\cdot \frac{\left(2\cdot\left(\frac{1}{y}\right)^{\frac{1}{2\beta+1}}\right)^{2\beta+1}}{2\beta+1}\right\} = \frac{2^{\beta+1}}{2\beta+1} < \infty. \qquad \square$$

Corollary 3.10. *Let a compact matrix quantum group $\mathbb{G}$ be of Kac type and $\widehat{\mathbb{G}}$ have the rapid decay property with respect to the natural length function on* Irr($\mathbb{G}$) *and with inequality* (3-3)*. Then, for each* $1 < p \leq 2$, *there exists a universal constant* $K = K(p) > 0$ *such that*

$$\left(\sum_{k\geq 0} \sum_{\substack{\alpha\in\mathrm{Irr}(\mathbb{G})\\ |\alpha|=k}} \frac{1}{(1+|\alpha|)^{(2-p)(\beta+1)}\left(\sum_{\beta\in S_k} n_\beta^2\right)^{\frac{2-p}{2}}} n_\alpha \|\hat{f}(\alpha)\|_{S_{n_\alpha}^p}^p\right)^{\frac{1}{p}} \leq K\|f\|_{L^p(\mathbb{G})} \tag{3-9}$$

for all $f \sim \sum_{\alpha\in\mathrm{Irr}(\mathbb{G})} n_\alpha \operatorname{tr}(\hat{f}(\alpha)u^\alpha) \in L^p(\mathbb{G})$.

Proof. Since $\frac{1}{p}=\frac{1}{2}+\frac{2-p}{2p}$ and $n_\alpha^{-\frac{1}{p}}\|\hat{f}(\alpha)\|_{S^p_{n_\alpha}}\le n_\alpha^{-\frac{1}{2}}\|\hat{f}(\alpha)\|_{\mathrm{HS}}$, we have

$$\sum_{\alpha\in S_k} n_\alpha\|\hat{f}(\alpha)\|^p_{S^p_{n_\alpha}}\le\sum_{\alpha\in S_k} n_\alpha^{2-\frac{p}{2}}\|\hat{f}(\alpha)\|^p_{\mathrm{HS}}=\big\|\big(n_\alpha^{\frac{1}{2}}\|\hat{f}(\alpha)\|_{\mathrm{HS}}\cdot n_\alpha^{\frac{2}{p}-1}\big)_{\alpha\in S_k}\big\|^p_{\ell^p(S_k)}$$

$$\le\big\|\big(n_\alpha^{\frac{1}{2}}\|\hat{f}(\alpha)\|_{\mathrm{HS}}\big)_{\alpha\in S_k}\big\|^p_{\ell^2(S_k)}\cdot\Big(\sum_{\alpha\in S_k}n_\alpha^2\Big)^{\frac{2-p}{2}}=\Big(\sum_{\alpha\in S_k}n_\alpha^2\Big)^{\frac{2-p}{2}}\cdot\Big(\sum_{\alpha\in S_k}n_\alpha\|\hat{f}(\alpha)\|^2_{\mathrm{HS}}\Big)^{\frac{p}{2}}.$$

Then we can obtain the conclusion above. □

4. Hardy–Littlewood inequalities

This section is dedicated to establishing explicit Hardy–Littlewood inequalities for the main targets: the reduced group C^*-algebras $C^*_r(G)$ with finitely generated discrete groups G, the free orthogonal quantum groups O_N^+ and the free permutation quantum groups S_N^+.

4A. *The reduced group C^*-algebras $C^*_r(G)$.* In this subsection, we deal with finitely generated discrete groups G. As expected, we find clear evidence that the geometric information of the underlying group is important for understanding noncommutative L^p-spaces $L^p(\mathrm{VN}(G))$.

Definition 4.1. A discrete group with a fixed finite symmetric generating set S is said to be polynomially growing if there exist $C>0$ and $k>0$ such that

$$\#\{g\in G:|g|\le n\}\le C(1+n)^k\quad\text{for all } n\ge 0.$$

In this case, the polynomial growth rate k_0 is defined as the minimum of such k. Then k_0 becomes a natural number and independent of the choice of generating set S.

Theorem 4.2. (1) *Let G be a finitely generated discrete group which has the polynomial growth rate k_0. Then, for each $1<p\le 2$, there exists a universal constant $K=K(p)$ such that*

$$\Big(\sum_{g\in G}\frac{1}{(1+|g|)^{(2-p)k_0}}|f(g)|^p\Big)^{\frac{1}{p}}\le K\|\lambda(f)\|_{L^p(\mathrm{VN}(G))}\tag{4-1}$$

for all $\lambda(f)\sim\sum_{g\in G}f(g)\lambda_g\in L^p(\mathrm{VN}(G))$.

(2) *Let G be a finitely generated discrete group with*

$$b_k=\#\{g\in G:|g|\le k\}\le Cr^k\quad\textit{for all } k\ge 0,$$

where $|\cdot|$ is the natural length function with respect to a finite symmetric generating set S. Then, for each $1<p\le 2$, there exists a universal constant $K=K(p,S)>0$ such that

$$\Big(\sum_{g\in G}\frac{1}{r^{(2-p)|g|}}|f(g)|^p\Big)^{\frac{1}{p}}\le K\|\lambda(f)\|_{L^p(\mathrm{VN}(G))}\tag{4-2}$$

for all $\lambda(f)\sim\sum_{g\in G}f(g)\lambda_g\in L^p(\mathrm{VN}(G))$.

Proof. (1) This is clear from Corollary 3.5.

(2) Consider $w(g) := 1/r^{|g|}$. Then

$$\sup_{t>0}\Big\{t\cdot \sum_{|g|\leq \log_r\left(\frac{1}{t}\right)} 1\Big\} = \sup_{0<t\leq 1}\Big\{t\cdot \sum_{|g|\leq \log_r\left(\frac{1}{t}\right)} 1\Big\} \leq C.$$

Then the conclusion is derived from Theorem 3.1 □

Remark 4.3. (1) For every finitely generated discrete group, there exist $C, r > 0$ such that $b_k \leq Cr^k$ for all $k \geq 0$ by Fekete's subadditivity lemma. Therefore, Theorem 4.2 covers all finitely generated discrete groups.

(2) In fact, we shall see that (4-1) is sharp because of Theorem 6.6.

Although we can always find inequality (4-2) for every finitely generated discrete group, we can achieve a better result by adding more detailed geometric information of the underlying groups. Indeed, if we assume hyperbolicity of a group, then the inequality is considerably improved.

Theorem 4.4. *Let G be any nonelementary word hyperbolic group with $b_k \leq Cr^k$ for all $k \geq 0$ with respect to a finite symmetric generating set S. Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(S, p)$ such that*

$$\begin{aligned}\Big(\sum_{g\in G}\frac{1}{r^{\frac{(2-p)|g|}{2}}(1+|g|)^{4-2p}}|f(g)|^p\Big)^{\frac{1}{p}} &\leq \Big(\sum_{k\geq 0}\frac{1}{(k+1)^{4-2p}}\Big(\sum_{\substack{g\in G\\ |g|=k}}|f(g)|^2\Big)^{\frac{p}{2}}\Big)^{\frac{1}{p}}\\ &\leq K\|\lambda(f)\|_{L^p(\mathrm{VN}(G))} \qquad (4\text{-}3)\end{aligned}$$

for all $\lambda(f) \sim \sum_{g\in G} f(g)\lambda_g \in L^p(\mathrm{VN}(G))$.

Proof. The conclusion comes from Corollary 3.10 and [de la Harpe 1988]. □

4B. ***Free quantum groups.*** Let us begin with the investigation of "genuine" quantum examples: the free orthogonal quantum groups O_N^+ and the free permutation quantum groups S_{N+2}^+. Moreover, we will find a subclass of $L^p(O_N^+)$ where the Hardy–Littlewood inequalities (4-4) and (4-5) on O_N^+ become equivalence (4-7), as for the result of SU(2) [Akylzhanov et al. 2015, Theorem 2.10].

Theorem 4.5. (1) *Let $\mathbb{G}$ be the free orthogonal quantum group O_2^+ or the free permutation quantum group S_4^+. Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(p)$ such that*

$$\Big(\sum_{k\geq 0}\frac{1}{(1+k)^{6-3p}}n_k\|\hat{f}(k)\|_{S^p_{n_k}}^p\Big)^{\frac{1}{p}} \leq \Big(\sum_{k\geq 0}\frac{1}{(1+k)^{6-3p}}n_k^{2-\frac{p}{2}}\|\hat{f}(k)\|_{\mathrm{HS}}^p\Big)^{\frac{1}{p}} \leq K\|f\|_{L^p(\mathbb{G})} \qquad (4\text{-}4)$$

for all $f \sim \sum_{k\geq 0} n_k \operatorname{tr}(\hat{f}(k)u^k) \in L^p(\mathbb{G})$.

(2) *Let $\mathbb{G}$ be a free orthogonal quantum group O_N^+ or a free permutation quantum group S_{N+2}^+ with $N \geq 3$. Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(p)$ such that*

$$\begin{aligned}\Big(\sum_{k\geq 0} \frac{1}{r_0^{(2-p)k}(1+k)^{4-2p}} n_k \|\hat{f}(k)\|_{S^p_{n_k}}^p\Big)^{\frac{1}{p}} &\leq \Big(\sum_{k\geq 0} \frac{1}{r_0^{(2-p)k}(1+k)^{4-2p}} n_k^{2-\frac{p}{2}} \|\hat{f}(k)\|_{\mathrm{HS}}^p\Big)^{\frac{1}{p}} \\ &\leq K\|f\|_{L^p(\mathbb{G})} \end{aligned} \tag{4-5}$$

for all $f \sim \sum_{k\geq 0} n_k \operatorname{tr}(\hat{f}(k)u^k) \in L^p(\mathbb{G})$, where $r_0 = \frac{1}{2}(N + \sqrt{N^2-4})$.

Proof. (1) In this case, $n_k = k+1$ (resp. $2k+1$) for all k. Thus, the conclusion comes from Corollary 3.5.

(2) It is known that $\widehat{O}_N^+$ and $\widehat{S}_{N+2}^+$ with $N \geq 3$ have the rapid decay property with $\beta = 1$ [Vergnioux 2007; Brannan 2013]. Also, $s_k = n_k^2 \approx r_0^{2k}$ for all $k \in \{0\} \cup \mathbb{N}$. Therefore, Corollaries 3.9 and 3.10 complete the proof. □

Remark 4.6. All results of this paper for S_N^+ can be extended to quantum automorphism group $\mathbb{G}_{\mathrm{aut}}(B, \psi)$ with a δ-trace ψ and $\dim(B) = N$ by repeating the same proofs. See [Brannan 2012a; 2013].

An important observation for the free orthogonal quantum groups O_N^+ is that the inequalities (4-4) and (4-5) become equivalences (4-7) under several assumptions. Essentially, this is based on the result of SU(2) [Akylzhanov et al. 2015, Theorem 2.10] and the following lemma moves the result to O_N^+.

Lemma 4.7. *Let $\mathbb{G} = O_N^+$ or S_{N+2}^+ with $N \geq 2$ and consider $G = \mathrm{SU}(2)$ or $\mathrm{SO}(3)$ for each case. Then, for $f \sim \sum_{n\geq 0} c_n \chi_n^1 \in L^p(\mathbb{G})$, the associated function $\Phi(f) \sim \sum_{n\geq 0} c_n \chi_n^2 \in L^p(G)$ has the same norm. More precisely,*

$$\|f\|_{L^p(\mathbb{G})} = \Big\|\Phi(f) \sim \sum_{n\geq 0} c_n \chi_n^2\Big\|_{L^p(G)} \quad \text{for all } 1 \leq p \leq \infty.$$

Here, $\chi_n^1 = \operatorname{tr}(u^n)$ and $\chi_n^2 = \operatorname{tr}(v^n)$, where u^n and v^n are the n-th irreducible unitary representations of $\mathbb{G}$ and G, respectively.

Proof. In the cases above, it is known that $\mathbb{G}$ and G share the same fusion rule. In [Wang 2017, Proposition 6.7], it was pointed out that the restricted map $\Phi|_{\mathrm{Pol}(\mathbb{G})}$ is a trace-preserving $*$-isomorphism. Now, for any $x = \sum_{k\geq 0} c_k \chi_k^1 \in \mathrm{Pol}(\mathbb{G})$ and $m \in \mathbb{N}$,

$$\begin{aligned} h((x^*x)^m) &= h\bigg(\Big(\sum_{k,l\geq 0} \bar{c}_k c_l (\chi_k^1)^* \chi_l^1\Big)^m\bigg) \\ &= \sum_{k_1,l_1,\dots,k_m,l_m\geq 0} \overline{c_{k_1}\cdots c_{k_m}} c_{l_1}\cdots c_{l_m} h\big((\chi_{k_1}^1)^*\chi_{l_1}^1\cdots(\chi_{k_m}^1)^*\chi_{l_m}^1\big) \\ &= \sum_{k_1,l_1,\dots,k_m,l_m\geq 0} \overline{c_{k_1}\cdots c_{k_m}} c_{l_1}\cdots c_{l_m} \int_G \bar{\chi}_{k_1}^2 \chi_{l_1}^2 \cdots \bar{\chi}_{k_m}^2 \chi_{l_m}^2 \\ &= \int_G \Big(\sum_{k,l\geq 0} \bar{c}_k c_l \bar{\chi}_k^2 \chi_l^2\Big)^m = \int_G |x'|^{2m}, \end{aligned}$$

where $x' = \sum_{k\geq 0} c_k \chi_k^2 \in \mathrm{Pol}(G)$. Then the Stone–Weierstrass theorem completes the proof. □

Corollary 4.8. *Let $N \geq 2$, $\frac{3}{2} < p \leq 2$ and fix $D > 0$. Also, assume $f \sim \sum_{k\geq 0} c_k \chi_k \in L^{\frac{3}{2}}(O_N^+)$ satisfies*

$$c_k \geq c_{k+1} \geq 0 \quad \text{and} \quad \sum_{m\geq k} \frac{c_m}{m+1} \leq D \cdot c_k \quad \text{for } k \geq 0. \tag{4-6}$$

Then we have

$$\|f\|_{L^p(O_N^+)} \approx \Bigl(\sum_{k\geq 0} (1+k)^{2p-4} c_k^p\Bigr)^{\frac{1}{p}}. \tag{4-7}$$

Proof. It is sufficient to combine the Lemma 4.7 and [Akylzhanov et al. 2015, Theorem 2.10]. □

5. A strong Hardy–Littlewood inequality

The studies of Hardy–Littlewood inequalities in [Akylzhanov et al. 2015; Hardy and Littlewood 1927; Hewitt and Ross 1974] dealt with general L^p-functions, but plenty of classical results of harmonic analysis on $\mathbb{T}$ show that a theorem on a function space can have a stronger form when restricted to a holomorphic setting [Kemp and Speicher 2007].

Evidence of these phenomena in the noncommutative setting is the strong Haagerup inequality on the reduced group C^*-algebras $C_r^*(\mathbb{F}_N)$. More precisely, it was shown that the rapid decay property can be strengthened in a general holomorphic setting [Kemp and Speicher 2007]. Such a phenomenon also occurs on the free unitary quantum groups [Brannan 2012b; 2012a].

Let $g_1, \ldots, g_N$ be canonical generators of $\mathbb{F}_N$ and denote by $\mathbb{F}_N^+$ the set of elements of the form $g_{i_1} g_{i_2} \cdots g_{i_m}$ with $m \in \{0\} \cup \mathbb{N}$ and $1 \leq i_k \leq N$ for all $1 \leq k \leq m$.

Theorem 5.1 (strong Haagerup inequality on $C_r^*(\mathbb{F}_N)$). *Consider a subset $E := \mathbb{F}_N^+$ and $E_k := \{g \in E : |g| = k\}$. Then, for any $k \in \{0\} \cup \mathbb{N}$, we have*

$$\Bigl\| \sum_{g\in E_k} f(g)\lambda_g \Bigr\|_{C_r^*(\mathbb{F}_N)} \leq \sqrt{e}\sqrt{k+1} \Bigl(\sum_{g\in E_k} |f(g)|^2 \Bigr)^{\frac{1}{2}}.$$

Based on this information, we can modify the inequality (3-4) as follows.

Proposition 5.2. *Let $N \geq 2$. Then we have*

$$\|f\|_{A(\mathbb{F}_N)} \geq \frac{1}{\sqrt{e}} \sup_{k\geq 0} \frac{\bigl(\sum_{g\in E_k} |f(g)|^2\bigr)^{\frac{1}{2}}}{(k+1)^{\frac{1}{2}}} \quad \text{for all } f \in A(\mathbb{F}_N). \tag{5-1}$$

Proof. We can use the proof of Proposition 3.7 again in this case. The only difference is the improvement of $(1+k)^\beta$ to $(1+k)^{\frac{1}{2}}$ in inequality (3-5). Then we are able to reach a conclusion by restricting support of $x \in C_c(G)$ to $\mathbb{F}_N^+$ in the proof. □

Theorem 5.3. *Let $N \geq 2$. Then, for each $1 < p \leq 2$, there exists a universal constant $K = K(p) > 0$ such that*

$$\begin{aligned} \Bigl(\sum_{g\in\mathbb{F}_N} \frac{1}{(1+|g|)^{\frac{3}{2}(2-p)} N^{\frac{(2-p)|g|}{2}}} |f(g)|^p \Bigr)^{\frac{1}{p}} &\leq \Bigl(\sum_{k\geq 0} \frac{1}{(1+k)^{\frac{3}{2}(2-p)}} \Bigl(\sum_{\substack{g\in\mathbb{F}_N \\ |g|=k}} |f(g)|^2 \Bigr)^{\frac{p}{2}} \Bigr)^{\frac{1}{p}} \\ &\leq K \|\lambda(f)\|_{L^p(\mathrm{VN}(\mathbb{F}_N))} \end{aligned} \tag{5-2}$$

for all $\lambda(f) \sim \sum_{g\in\mathbb{F}_N} f(g)\lambda_g \in L^p(\mathrm{VN}(\mathbb{F}_N))$ with $\mathrm{supp}(f) \subseteq \mathbb{F}_N^+$.

Proof. It can be obtained by repeating the proofs of Theorem 3.8 and Corollary 3.10. The only difference is that it the operator A should be replaced with

$$\lambda(f) \mapsto \left(\frac{\| f \cdot \chi_{E_k} \|_{\ell^2(\mathbb{F}_N)}}{w(k)} \right)_{k \geq 0}$$

and $(1+k)^{\beta}$ with $(1+k)^{\frac{1}{2}}$. Also, we choose a weight function w on $\{0\} \cup \mathbb{N}$ by $w(k) := 1/(1+k)^{\frac{3}{2}}$. Then we can get a new inequality for general $\lambda(f) \in L^p(\mathrm{VN}(\mathbb{F}_N))$, but our consideration is in the case $\mathrm{supp}(f) \subseteq \mathbb{F}_N^+$. □

6. Sharpness

Hardy–Littlewood inequalities discussed in Section 4 give a specific pair (r, s) such that the multiplier

$$\mathcal{F}_{w_{r,s}} : L^p(\mathbb{G}) \to \ell^p(\widehat{\mathbb{G}}), \quad f \mapsto (w_{r,s}(\alpha) \hat{f}(\alpha))_{\alpha \in \mathrm{Irr}(\mathbb{G})},$$

is bounded for each case, where $w_{r,s}(\alpha) = 1/r^{|\alpha|}(1+|\alpha|)^s$ with respect to the natural length $|\cdot|$ on $\mathrm{Irr}(\mathbb{G})$.

Here is the list of such pairs:

- $\left(0, \frac{n(2-p)}{p}\right)$ for G a compact Lie group,
- $\left(0, \frac{k_0(2-p)}{p}\right)$ for $C_r^*(G)$ with a polynomially growing discrete group G,
- $\left(0, \frac{3(2-p)}{p}\right)$ for O_2^+ or S_4^+ and
- $\left(r_0^{\frac{2-p}{p}}, \frac{2(2-p)}{p}\right)$ for O_N^+ or S_{N+2}^+ with $N \geq 3$.

Remark 6.1 [Lee and Youn 2017; Wallach 1973]. If G is a compact Lie group, then $\sqrt{\kappa_\pi}$ is equivalent to the natural length function $\|\cdot\|_S$ generated by the fundamental generating set S of $\mathrm{Irr}(G)$. Equivalently, $(1+\kappa_\pi)^{\frac{\beta}{2}} \approx (1+\|\pi\|_S)^{\beta}$.

In order to claim that the established inequalities are sharp, we will show that there is no (r', s') better than the given specific pair in that $\mathcal{F}_{w_{r',s'}}$ is unbounded whenever (1) $r' < r$ or (2) $r' = r,\ s' < s$.

This viewpoint is different from the spirit of [Akylzhanov et al. 2015, Theorem 2.10] or Corollary 4.8, which requires finding an equivalence in a subclass. However, our approach is quite natural since it is strongly related to the Sobolev embedding theorem. In this section and Section 7B, we will discuss how they are related. For example, $\mathcal{F}_{w_{0,s}} : L^p(\mathbb{T}^d) \to \ell^p(\mathbb{Z}^d)$ is bounded if and only if $H_p^s(\mathbb{T}^d) \subseteq L^{p'}(\mathbb{T}^d)$ and it is equivalent to

$$H_q^{\frac{ps}{2-p}\left(\frac{1}{q}-\frac{1}{r}\right)}(\mathbb{T}^d) \subseteq L^r(\mathbb{T}^d)$$

for all $1 < q < r < \infty$, where $H_p^s(\mathbb{T}^d)$ is the Bessel potential space.

In addition, this view has a definite advantage over looking for equivalence because we can cover a much larger class.

Our first strategy is to handle an ultracontractivity problem on $C(G)$ with compact Lie groups and $C_r^*(G)$ with polynomially growing discrete groups. In fact, ultracontractivity problems are strongly related to Sobolev embedding properties [Xiong 2016].

Let M be the von Neumann subalgebra generated by $\{\chi_\alpha\}_{\alpha\in\mathrm{Irr}(\mathbb{G})}$ in $L^\infty(\mathbb{G})$ and consider $L^p(M)$, the noncommutative L^p-space associated with the restriction of the Haar state on M. Now suppose that $l:\mathrm{Irr}(\mathbb{G})\to(0,\infty)$ is a positive function and there exist $1<p<2$ and a universal constant $C>0$ such that

$$\Big\| J(f)\sim\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{1}{l(\alpha)^{\frac{\beta}{p}}}c_\alpha\chi_\alpha\Big\|_{L^{p'}(M)}\le C\Big\| f\sim\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}c_\alpha\chi_\alpha\Big\|_{L^p(M)}, \tag{6-1}$$

where J is a densely defined positive operator on $L^2(M)$ which maps χ_α to $1/l(\alpha)^{\frac{\beta}{p}}\chi_\alpha$ for all $\alpha\in\mathrm{Irr}(\mathbb{G})$, $1\le i,j\le n_\alpha$. Indeed, $J=K^*K$, where $K:\chi_\alpha\mapsto 1/l(\alpha)^{\frac{\beta}{2p}}\chi_\alpha$.

Now take $\phi(t):=t^{\frac{2\beta}{2-p}}$, $\psi(z):=z^{\frac{2\beta}{2-p}}$ and $L:=J^{-\frac{p}{2\beta}}$. Then [Xiong 2016, Theorem 1.1] suggests that there exists a universal constant $C'>0$ such that

$$\Big\| e^{-tL}(f)\sim\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{c_\alpha}{e^{tl(\alpha)^{1/2}}}\chi_\alpha\Big\|_{L^\infty(M)}\le C'\frac{\|f\|_{L^1(M)}}{t^{\frac{2\beta}{2-p}}} \tag{6-2}$$

for all $0<t<\infty$ and all $f\sim\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}c_\alpha\chi_\alpha\in L^1(M)$.

The next thing to do is to prove that the following result (6-3) is achieved by combining (6-2), Lemma 6.2 and Lemma 6.3 below:

$$\sup_{0<t<\infty}\Big\{t^{\frac{2\beta}{2-p}}\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{e^{tl(\alpha)^{1/2}}}\Big\}=:C<\infty, \tag{6-3}$$

if $\mathbb{G}$ is G a compact Lie group or $C_r^*(G)$ with a polynomially growing discrete group.

Lemma 6.2 [Dasgupta and Ruzhansky 2014, Lemma 4.1; Lee and Youn 2017, Proposition 5.7].

(1) *Let G be a compact Lie group with the dimension n. Then*

$$\sum_{\pi\in\mathrm{Irr}(G)}\frac{n_\pi^2}{(1+\kappa_\pi)^{\frac{s}{2}}}<\infty$$

if and only if $s>n$.

(2) *Let G be a finitely generated discrete group with the polynomial growth rate k_0. Then*

$$\sum_{g\in G}\frac{1}{(1+|g|)^s}<\infty$$

if and only if $s>k_0$.

Lemma 6.3. (1) *Let G be a compact Lie group. Then there exist probability measures $\{\nu_t\}_{t>0}$ such that $\hat{\nu}_t(\pi)=1/e^{t\kappa_\pi}\mathrm{Id}_{n_\pi}$ for all $\pi\in\mathrm{Irr}(G)$. Moreover, $\{\nu_t\}_{t>0}\subseteq L^1(G)$.*

(2) *Let G be a compact Lie group and let $f\sim\sum_{\pi\in\mathrm{Irr}(G)}n_\pi\,\mathrm{tr}(\hat{f}(\pi)\pi)\in L^\infty(G)$ be such that $\hat{f}(\pi)\ge 0$ for all $\pi\in\mathrm{Irr}(G)$. Then*

$$\|f\|_{L^\infty(G)}=\sum_{\pi\in\mathrm{Irr}(G)}n_\pi\,\mathrm{tr}(\hat{f}(\pi)).$$

(3) *Let G be a discrete group. Then the Fourier algebra $A(G)$ has a bounded approximate identity if and only if G is amenable. In this case, the bounded approximate identity can be chosen as positive and compactly supported functions on G.*

(4) *If G is an amenable discrete group, we have*

$$\Big\|\lambda(f) \sim \sum_{g\in G} f(g)\lambda_g\Big\|_{\mathrm{VN}(G)} = \sum_{g\in G} f(g)$$

for any positive function $f \in \ell^1(G)$.

Proof. (1) Since

$$\sum_{\pi\in \mathrm{Irr}(G)} \frac{n_\pi^2}{e^{t\kappa_\pi}} < \infty$$

by Lemma 6.2, we know that $\nu_t \in A(G) \subseteq C(G) \subseteq L^1(G)$. The family $\{\nu_t\}_{t>0}$ is called the heat semigroup of measures.

(2) Since $f \mapsto \mu_t * f$ is a contractive map on $L^\infty(G)$ for all $t > 0$, where $*$ is the convolution product, we have

$$\|f\|_{L^\infty(G)} \geq \sup_{t>0}\Big\| f_t \sim \sum_{\pi\in \mathrm{Irr}(G)} \frac{n_\pi}{e^{t\kappa_\pi}} \operatorname{tr}(\hat{f}(\pi)\pi)\Big\|_{C(G)}.$$

Here, since

$$\sum_\pi \frac{n_\pi}{e^{t\kappa_\pi}} \operatorname{tr}(\hat{f}(\pi)) \leq \|f\|_{L^1(G)} \sum_\pi \frac{n_\pi^2}{e^{t\kappa_\pi}} < \infty$$

by Lemma 6.2, the Fourier series of f_t uniformly converges to $f_t \in C(G)$. Therefore,

$$\|f\|_{L^\infty(G)} \geq \sup_{t>0} f_t(1) = \sup_{t>0} \sum_{\pi\in \mathrm{Irr}(G)} \frac{n_\pi}{e^{t\kappa_\pi}} \operatorname{tr}(\hat{f}(\pi)) = \sum_{\pi\in \mathrm{Irr}(G)} n_\pi \operatorname{tr}(\hat{f}(\pi)).$$

The other direction is trivial.

(3) See [Runde 2002, Theorem 7.1.3] and its proof. We may also assume the compact supportness by considering $f_\epsilon := f \cdot \chi_{\{g\in G: f(g)>\epsilon\}}$ for positive $f \in \ell^1(G)$.

(4) This is Kesten's condition, which is equivalent to amenability. □

Now we can prove that the claim is true.

Proposition 6.4. *Let $\mathbb{G}$ be G a compact Lie group or $C_r^*(G)$ with a polynomially growing discrete group. Also, suppose that the inequality* (6-1) *holds. Then*

$$\sup_{0<t<\infty}\left\{ t^{\frac{2\beta}{2-p}} \sum_{\alpha\in \mathrm{Irr}(\mathbb{G})} \frac{n_\alpha^2}{e^{tl(\alpha)^{1/2}}}\right\} =: C < \infty. \tag{6-4}$$

Proof. For G a compact Lie group, by Lemma 6.3, for all $0 < t < \infty$,

$$\sum_{\pi\in \mathrm{Irr}(G)} \frac{n_\pi^2}{e^{tl(\pi)^{1/2}}} = \sup_{r>0} \sum_{\pi\in \mathrm{Irr}(G)} \frac{n_\pi^2}{e^{tl(\pi)^{1/2}} e^{r\kappa_\pi}} = \sup_{r>0} \|e^{-tL}(\nu_r)\|_{L^\infty(G)} \leq \frac{C'}{t^{\frac{2\beta}{2-p}}} \sup_{r>0} \|\nu_r\|_{L^1(G)} \leq \frac{C'}{t^{\frac{2\beta}{2-p}}}.$$

Now, for G a polynomially growing discrete group, there exists a bounded approximate identity $(e_i)_i$ in $A(G)$ that consists of positive and compactly supported functions since polynomially growing discrete groups are always amenable. Then (6-2) implies

$$\sum_{g\in G}\frac{1}{e^{tl(g)^{1/2}}}=\sup_i\sum_{g\in G}\frac{e_i(g)}{e^{tl(g)^{1/2}}}=\sup_i\left\|\sum_{g\in G}\frac{e_i(g)}{e^{tl(g)^{1/2}}}\lambda_g\right\|_{C_r^*(G)}\le\frac{C''}{t^{\frac{2\beta}{2-p}}}$$

since $\lim_i e_i(g)=1$ for all $g\in G$. □

Proposition 6.4 allows us to extract a quantitative observation.

Proposition 6.5. *Let $\mathbb{G}$ be G a compact Lie group or $C_r^*(G)$ with a polynomially growing discrete group. Also, suppose that the inequality* (6-1) *holds. Then we have*

$$\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{l(\alpha)^{\frac{m}{2}}}<\infty\quad\textit{for all natural numbers } m>\frac{2\beta}{2-p}. \tag{6-5}$$

Proof. Choose $\gamma\in\left(\max\left\{\frac{2\beta}{2-p},m-1\right\},m\right)$. Then we have

$$\sup_{0<t\le 1}\left\{t^\gamma\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{e^{tl(\alpha)^{1/2}}}\right\}=:C_0<\infty$$

from (6-4), so that

$$\int_t^1\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{e^{xl(\alpha)^{1/2}}}\,dx\le C_0\int_t^1\frac{1}{x^\gamma}\,dx$$

for all $0<t\le 1$.

This implies

$$\sup_{0<t\le 1}\left\{t^{\gamma-1}\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{l(\alpha)^{\frac{1}{2}}e^{tl(\alpha)^{1/2}}}\right\}=:C_1<\infty,$$

so that we can inductively see that

$$\sup_{0<t\le 1}\left\{t^{\gamma-(m-1)}\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{l(\alpha)^{\frac{m-1}{2}}e^{tl(\alpha)^{1/2}}}\right\}=:C_{m-1}<\infty.$$

Then there exist $D_1,D_2>0$ such that

$$\sum_{\alpha\in\mathrm{Irr}(\mathbb{G})}\frac{n_\alpha^2}{l(\alpha)^{\frac{m}{2}}e^{tl(\alpha)^{1/2}}}\le D_1t^{m-\gamma}+D_2\quad\text{for all } 0<t\le 1$$

through similar reasoning.

Lastly, taking the limit $t\to 0$ completes the proof. □

Theorem 6.6. *Let $1<p\le 2$.*

(1) *Let G be a compact Lie group with dimension n. Then*

$$\left(\sum_{\pi\in\mathrm{Irr}(G)}\frac{1}{(1+\kappa_\pi)^{\frac{s}{2}}}n_\pi\|\hat{f}(\pi)\|_{S_{n_\pi}^p}^p\right)^{\frac{1}{p}}\lesssim\|f\|_{L^p(G)}$$

holds if and only if $s\ge n(2-p)$.

(2) *Let G be a finitely generated discrete group with polynomial growth rate k_0. Then*

$$\left(\sum_{g\in G}\frac{1}{(1+|g|)^s}|f(g)|^p\right)^{\frac{1}{p}} \lesssim \|\lambda(f)\|_{L^p(\mathrm{VN}(G))}$$

holds if and only if $s \geq k_0(2-p)$.

(3) *Let $\mathbb{G}$ be O_2^+ or S_4^+. Then*

$$\left(\sum_{k\geq 0}\frac{1}{(1+k)^s}n_k\|\hat{f}(k)\|^p_{S^p_{n_k}}\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})}$$

holds if and only if $s \geq 3(2-p)$.

(4) *Let $\mathbb{G}$ be O_N^+ or S_{N+2}^+ with $N \geq 3$. Then*

$$\left(\sum_{k\geq 0}\frac{1}{r_0^{(2-p)k}(1+k)^s}n_k\|\hat{f}(k)\|^p_{S^p_{n_k}}\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})}$$

holds if and only if $s \geq 4-2p$, where $r_0 = \frac{1}{2}(N+\sqrt{N^2-4})$.

Proof. One direction is obtained from the discussed Hardy–Littlewood inequalities (1-3), (4-1), (4-4) and (4-5). To prove the converse direction, firstly, define $l(\alpha)$ by (1) $(1+\kappa_\pi)^{\frac{1}{2}}$ and (2) $1+|g|$ respectively. Then the assumed inequality

$$\left(\sum_{\alpha\in \mathrm{Irr}(\mathbb{G})}\frac{n_\alpha}{l(\alpha)^s}\|\hat{f}(\alpha)\|^p_{S^p_{n_\alpha}}\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})}$$

implies the inequality (6-1) for $\beta = s$. Then, by Proposition 6.5 and Lemma 6.2, we can get (1) $2n \leq \frac{2\beta}{2-p}$, (2) $2k_0 \leq \frac{2\beta}{2-p}$ respectively.

In (3) and (4), let $G = \mathrm{SU}(2)$ if $\mathbb{G} = O_N^+$ and $G = \mathrm{SO}(3)$ if $\mathbb{G} = S_{N+2}^+$ for each case. Also, denote by $\chi_k' \in \mathrm{Pol}(G)$ the character corresponding to $\chi_k \in \mathrm{Pol}(\mathbb{G})$, as in Lemma 4.7. Define $l(k) := 1+k$.

First of all, in (3), for each $f \sim \sum_{k\geq 0} c_k\chi_k \in L^p(\mathbb{G})$, we have

$$\left\|\sum_{k\geq 0}(1+k)^{-\frac{s}{p}}c_k\chi_k\right\|_{L^{p'}(G)} \leq \left(\sum_{k\geq 0}\frac{1}{(1+k)^{s+p-2}}|c_k|^p\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})} = \left\|f' \sim \sum_{k\geq 0}c_k\chi_k'\right\|_{L^p(G)}$$

by Lemma 4.7 and the Hausdorff–Young inequality. On the other hand, in (4), for each $f \sim \sum_{k\geq 0} c_k\chi_k \in L^p(\mathbb{G})$, we have

$$\left\|\sum_{k\geq 0}(1+k)^{-\frac{s+2-p}{p}}c_k\chi_k\right\|_{L^{p'}(G)} \leq \left(\sum_{k\geq 0}\frac{1}{(1+k)^{s}}|c_k|^p\right)^{\frac{1}{p}} \lesssim \|f\|_{L^p(\mathbb{G})} = \left\|f' \sim \sum_{k\geq 0}c_k\chi_k'\right\|_{L^p(G)}$$

by similar arguments.

Now we can apply Proposition 6.5 and Lemma 6.2 for compact Lie groups again, so that (3) $s \geq 6-3p$ and (4) $s-p+2 \geq 6-3p (\Leftrightarrow s \geq 4-2p)$ respectively. $\square$

7. Some remarks about Sidon sets, the Sobolev embedding theorem and quantum torus

The methods of this paper, combined with a lacunarity result for compact quantum groups, give a Sobolev-embedding-theorem-type interpretation for G compact Lie groups and for $C_r^*(G)$ with polynomially growing groups. Also, we show an explicit inequality on the quantum torus $\mathbb{T}_\theta^d$.

7A. ***Sidon sets on compact quantum groups.*** The study of lacunarity, particularly Sidon sets, is one of the major subjects in harmonic analysis and recently the notion has been extended to the setting of compact quantum groups [Wang 2017].

Definition 7.1. Let $\mathbb{G}$ be a compact quantum group.

(1) A subset $E \subseteq \mathrm{Irr}(\mathbb{G})$ is called a Sidon set if there exists $K > 0$ such that

$$\|\hat{f}\|_{\ell^1(\widehat{\mathbb{G}})} \leq K\|f\|_{L^\infty(\mathbb{G})} \quad \text{for all } f \in \mathrm{Pol}_E(\mathbb{G}),$$

where $\mathrm{Pol}_E(\mathbb{G}) := \{f \in \mathrm{Pol}(\mathbb{G}) : \hat{f}(\alpha) = 0 \text{ for all } \alpha \notin E\}$.

(2) A subset $E \subseteq \mathrm{Irr}(\mathbb{G})$ is called a central Sidon set if there exists $K > 0$ such that

$$\|\hat{f}\|_{\ell^1(\widehat{\mathbb{G}})} \leq K\|f\|_{L^\infty(\mathbb{G})} \quad \text{for all } f = \sum_{\alpha \in \mathrm{Irr}(\mathbb{G})} c_\alpha \chi_\alpha \in \mathrm{Pol}_E(\mathbb{G}).$$

Let $\mathbb{G} = (A, \Delta)$ be of Kac type and $E \subseteq \mathrm{Irr}(\mathbb{G})$ be a central Sidon set. Then [Wang 2017, Proposition 6.4] implies there exists $\mu \in M(\mathbb{G}) = C_r(\mathbb{G})^*$ such that $\hat{\mu}(\alpha) = (\mu((u_{j,i}^\alpha)^*))_{1 \leq i,j \leq n_\alpha} = \mathrm{Id}_{n_\alpha}$ for all $\alpha \in E$. Since $\mathrm{Pol}(\mathbb{G})$ is dense in $C_r(\mathbb{G})$, Proposition 3.7 still holds for $\mu \in M(\mathbb{G})$.

Now, if $\mathbb{G}$ satisfies the assumptions of Proposition 3.7 and if $E \subseteq \mathrm{Irr}(\mathbb{G})$ is a central Sidon set, we get

$$\infty > \sup_{k \geq 0} \frac{\left(\sum_{\alpha \in E_k} n_\alpha^2\right)^{\frac{1}{2}}}{(1+k)^\beta} \geq \sup_{k \geq 0} \frac{|E_k|^{\frac{1}{2}} \min_{\alpha \in E_k} n_\alpha}{(1+k)^\beta},$$

where $E_k := \{\alpha \in E : |\alpha| = k\}$.

Thus, the conditions $|E| = \infty$ and $n_\alpha > r^{|\alpha|}$ for all $\alpha \in \mathrm{Irr}(\mathbb{G})$ with $r > 1$ cannot hold true simultaneously.

Remark 7.2. (1) The argument above shows that there is no infinite (central) Sidon set in U_N^+ with $N \geq 3$, which is not explained in [Wang 2017].

(2) Shortly after this research, the author of [Wang 2017] personally informed me of a simple idea to further explain the U_N^+ cases. Under the identification $\mathrm{Irr}(U_N^+) \cong \mathbb{F}_2^+$, the fact that

$$\|\chi_\alpha\|_4 = (1+|\alpha|)^{\frac{1}{4}} \quad \text{for all } \alpha \in \mathbb{F}_2^+$$

implies there is no infinite $\Lambda(4)$ set, so that there is no infinite Sidon set on U_N^+ with $N \geq 2$.

7B. ***Sobolev embedding properties.*** The content of Section 6 can be interpreted in terms of Sobolev embedding properties by [Xiong 2016, Theorem 1.1].

For G a compact Lie group whose real dimension is n, the computations in Section 6 suggest that

$$\Bigl\|(1-\Delta)^{-\frac{\beta}{2}}(f)\sim\sum_{\pi\in\mathrm{Irr}(G)}\frac{n_\pi}{(1+\kappa_\pi)^{\frac{\beta}{2}}}\operatorname{tr}(\hat{f}(\pi)\pi)\Bigr\|_{L^{p'}(G)}\lesssim\|f\|_{L^p(G)}$$

if and only if $\beta\geq\frac{n(2-p)}{p}$ for each $1<p\leq 2$. Moreover, it is equivalent to

$$\bigl\|(1-\Delta)^{-\frac{\beta}{2}(\frac{1}{p}-\frac{1}{q})}(f)\bigr\|_{L^q(G)}\lesssim\|f\|_{L^p(G)}$$

if and only if $\beta\geq n$ for each $1<p<q<\infty$. If we define the space $H_p^s(G):=\{f\in L^p(G):(1-\Delta)^{\frac{s}{2}}(f)\in L^p(G)\}$ as an analogue of the Bessel potential space, then the result above is interpreted as

$$H_p^s(G)\subseteq L^q(G)\quad\text{if and only if}\quad s\geq n\bigl(\tfrac{1}{p}-\tfrac{1}{q}\bigr)\tag{7-1}$$

for each $1<p<q<\infty$.

On the other hand, if G is a finitely generated discrete group with polynomial growth rate k_0, then we can define an infinitesimal generator L on $C_r^*(G)$ by $\lambda_g\mapsto -|g|\lambda_g$ for all $g\in G$. Then we are able to induce the Sobolev embedding property of noncommutative spaces $L^p(\mathrm{VN}(G))$ as follows:

$$\bigl\|(1-L)^{-\beta(\frac{1}{p}-\frac{1}{q})}(\lambda(f))\bigr\|_{L^q(\mathrm{VN}(G))}\lesssim\|\lambda(f)\|_{L^p(\mathrm{VN}(G))}\quad\text{if and only if}\quad\beta\geq k_0\tag{7-2}$$

for each $1<p<q<\infty$.

The reader may consider another natural infinitesimal generator $L':\lambda_g\mapsto -|g|^2\lambda_g$, but the result is essentially the same when $(1-L)$ is replaced with $(1-L')^{\frac{1}{2}}$.

7C. *Hardy–Littlewood inequality on quantum torus.* The quantum torus $\mathbb{T}_\theta^d$ is a widely studied example of quantum space, but it is not a quantum group [Sołtan 2010]. Nevertheless, we can establish Hardy–Littlewood inequalities on $\mathbb{T}_\theta^d$, which is of the same form as the case for $\mathbb{T}^d$. A proof can be given by repeating the proof of Theorem 3.1. See [Xiong et al. 2015] for Fourier analysis on the quantum torus.

Remark 7.3. For a quantum torus $\mathbb{T}_\theta^d$, for each $1<p\leq 2$, we have

$$\biggl(\sum_{m\in\mathbb{Z}^d}\frac{1}{(1+\|m\|_1)^{d(2-p)}}|\hat{x}(m)|^p\biggr)^{\frac{1}{p}}\lesssim\|x\|_{L_p(\mathbb{T}_\theta^d)}.\tag{7-3}$$

Acknowledgements

The author is grateful to Hun Hee Lee for helpful comments on this work and would also like to thank Simeng Wang for discussing the Sidon set of compact quantum groups.

References

[Akylzhanov et al. 2015] R. Akylzhanov, E. Nursultanov, and M. Ruzhansky, "Hardy–Littlewood, Hausdorff–Young–Paley inequalities, and L^p-L^q Fourier multipliers on compact homogeneous manifolds", preprint, 2015. arXiv

[Akylzhanov et al. 2016] R. Akylzhanov, E. Nursultanov, and M. Ruzhansky, "Hardy–Littlewood–Paley inequalities and Fourier multipliers on SU(2)", *Studia Math.* **234**:1 (2016), 1–29. MR Zbl

[Banica and Vergnioux 2009] T. Banica and R. Vergnioux, "Growth estimates for discrete quantum groups", *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **12**:2 (2009), 321–340. MR Zbl

[Bekjan and Chen 2012] T. N. Bekjan and Z. Chen, "Interpolation and Φ-moment inequalities of noncommutative martingales", *Probab. Theory Related Fields* **152**:1-2 (2012), 179–206. MR Zbl

[Bényi and Oh 2013] A. Bényi and T. Oh, "The Sobolev inequality on the torus revisited", *Publ. Math. Debrecen* **83**:3 (2013), 359–374. MR Zbl

[Brannan 2012a] M. Brannan, *On the reduced operator algebras of free quantum groups*, Ph.D. thesis, Queen's University, 2012, available at https://search.proquest.com/docview/1512650272. MR

[Brannan 2012b] M. Brannan, "Quantum symmetries and strong Haagerup inequalities", *Comm. Math. Phys.* **311**:1 (2012), 21–53. MR Zbl

[Brannan 2013] M. Brannan, "Reduced operator algebras of trace-perserving quantum automorphism groups", *Doc. Math.* **18** (2013), 1349–1402. MR Zbl

[Caspers 2013] M. Caspers, "The L^p-Fourier transform on locally compact quantum groups", *J. Operator Theory* **69**:1 (2013), 161–193. MR Zbl

[Dasgupta and Ruzhansky 2014] A. Dasgupta and M. Ruzhansky, "Gevrey functions and ultradistributions on compact Lie groups and homogeneous spaces", *Bull. Sci. Math.* **138**:6 (2014), 756–782. MR Zbl

[Folland 1999] G. B. Folland, *Real analysis: modern techniques and their applications*, 2nd ed., John Wiley & Sons, New York, 1999. MR Zbl

[Franz et al. 2017] U. Franz, G. Hong, F. Lemeux, M. Ulrich, and H. Zhang, "Hypercontractivity of heat semigroups on free quantum groups", *J. Operator Theory* **77**:1 (2017), 61–76. MR

[Haagerup 1979] U. Haagerup, "L^p-spaces associated with an arbitrary von Neumann algebra", pp. 175–184 in *Algèbres d'opérateurs et leurs applications en physique mathématique* (Marseille, 1977), Colloq. Internat. CNRS **274**, CNRS, Paris, 1979. MR Zbl

[Hardy and Littlewood 1927] G. H. Hardy and J. E. Littlewood, "Some new properties of Fourier constants", *Math. Ann.* **97**:1 (1927), 159–209. MR

[de la Harpe 1988] P. de la Harpe, "Groupes hyperboliques, algèbres d'opérateurs et un théorème de Jolissaint", *C. R. Acad. Sci. Paris Sér. I Math.* **307**:14 (1988), 771–774. MR Zbl

[Hewitt and Ross 1974] E. Hewitt and K. A. Ross, "Rearrangements of L^r Fourier series on compact abelian groups", *Proc. London Math. Soc.* (3) **29** (1974), 317–330. MR Zbl

[Junge et al. 2014] M. Junge, T. Mei, and J. Parcet, "Smooth Fourier multipliers on group von Neumann algebras", *Geom. Funct. Anal.* **24**:6 (2014), 1913–1980. MR Zbl

[Junge et al. 2017] M. Junge, C. Palazuelos, J. Parcet, and M. Perrin, *Hypercontractivity in group von Neumann algebras*, Memoirs of the American Mathematical Society **1183**, Amer. Math. Soc., Providence, RI, 2017.

[Kemp and Speicher 2007] T. Kemp and R. Speicher, "Strong Haagerup inequalities for free $\mathscr{R}$-diagonal elements", *J. Funct. Anal.* **251**:1 (2007), 141–173. MR Zbl

[Kustermans and Vaes 2000] J. Kustermans and S. Vaes, "Locally compact quantum groups", *Ann. Sci. École Norm. Sup.* (4) **33**:6 (2000), 837–934. MR Zbl

[Kustermans and Vaes 2003] J. Kustermans and S. Vaes, "Locally compact quantum groups in the von Neumann algebraic setting", *Math. Scand.* **92**:1 (2003), 68–92. MR Zbl

[Lee and Youn 2017] H. H. Lee and S.-g. Youn, "New deformations of convolution algebras and Fourier algebras on locally compact groups", *Canad. J. Math.* **69**:2 (2017), 434–452. MR Zbl

[Pisier and Xu 2003] G. Pisier and Q. Xu, "Non-commutative L^p-spaces", pp. 1459–1517 in *Handbook of the geometry of Banach spaces, II*, edited by W. B. Johnson and J. Lindenstrauss, North-Holland, Amsterdam, 2003. MR Zbl

[Runde 2002] V. Runde, *Lectures on amenability*, Lecture Notes in Mathematics **1774**, Springer, 2002. MR Zbl

[Sołtan 2010] P. M. Sołtan, "When is a quantum space not a group?", pp. 353–364 in *Banach algebras 2009*, edited by R. J. Loy et al., Banach Center Publ. **91**, Polish Acad. Sci. Inst. Math., Warsaw, 2010. MR Zbl

[Timmermann 2008] T. Timmermann, *An invitation to quantum groups and duality: from Hopf algebras to multiplicative unitaries and beyond*, Eur. Math. Soc., Zürich, 2008. MR Zbl

[Vergnioux 2007] R. Vergnioux, "The property of rapid decay for discrete quantum groups", *J. Operator Theory* **57**:2 (2007), 303–324. MR Zbl

[Wallach 1973] N. R. Wallach, *Harmonic analysis on homogeneous spaces*, Pure and Applied Mathematics **19**, Marcel Dekker, New York, 1973. MR Zbl

[Wang 1995] S. Wang, "Free products of compact quantum groups", *Comm. Math. Phys.* **167**:3 (1995), 671–692. MR Zbl

[Wang 1998] S. Wang, "Quantum symmetry groups of finite spaces", *Comm. Math. Phys.* **195**:1 (1998), 195–211. MR Zbl

[Wang 2017] S. Wang, "Lacunary Fourier series for compact quantum groups", *Comm. Math. Phys.* **349**:3 (2017), 895–945. MR Zbl

[Woronowicz 1987] S. L. Woronowicz, "Twisted SU(2) group: an example of a noncommutative differential calculus", *Publ. Res. Inst. Math. Sci.* **23**:1 (1987), 117–181. MR Zbl

[Xiong 2016] X. Xiong, "Noncommutative harmonic analysis on semigroup and ultracontractivity", preprint, 2016. To appear in *Indiana Univ. Math. J.* arXiv

[Xiong et al. 2015] X. Xiong, Q. Xu, and Z. Yin, "Sobolev, Besov and Triebel–Lizorkin spaces on quantum tori", preprint, 2015. To appear in *Mem. Amer. Math. Soc.* arXiv

[Xu 2007] Q. Xu, "Operator spaces and noncommutative L_p: the part on noncommutative L_p-spaces", lecture notes, Nankai University, 2007, available at http://www.cim.nankai.edu.cn/activites/conferences/hy200707/summer-school/Xu-long.pdf.

Received 16 Feb 2017. Revised 16 Jun 2017. Accepted 24 Jul 2017.

SANG-GYUN YOUN: yun87654@snu.ac.kr
Department of Mathematical Sciences, Seoul National University, Seoul, South Korea

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 11 No. 1 2018

2157-5045(2018)11:1;1-T