Pacific Journal of Mathematics

Volume 291 No. 1 November 2017

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

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

The subscription price for 2017 is US $450/year for the electronic version, and $625/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.1

CHAIN TRANSITIVE HOMEOMORPHISMS ON A SPACE: ALL OR NONE

ETHAN AKIN AND JUHO RAUTIO

Extending earlier work, we consider when a compact metric space can be realized as the omega limit set of a discrete time dynamical system. This is equivalent to asking when the space admits a chain transitive homeomorphism. We approach this problem in terms of various conditions on the connected components of the space. We also construct spaces where all homeomorphisms are chain transitive.

1. Introduction

Since long-term behavior is a central concern in dynamical systems theory, it is natural to consider the set of limit points for the orbit of a point in the state space, obtained as time tends to infinity. This omega limit set was explicitly defined for a real flow by Birkhoff [1927] and appears as well in the classic book [Andronov and Khaikin 1937]. The discrete time version, i.e., for systems obtained by iterating a homeomorphism, was studied in detail by Dowker [1953] and with extensions to maps by Sharkovsky [1965].

We will use the term *space* to mean a nonempty, compact, metrizable space unless otherwise mentioned. When a metric is required, we assume that one is chosen and fixed. The results are independent of the choice of metric. Broadly, our metric space ideas and constructions are really uniform space concepts, and a compact space has a unique uniformity.

We will use $[a, b]$, (a, b), $[a, b)$, etc. to denote intervals in $\mathbb{R}$, and so we will use $\langle a, b\rangle$ to represent points of $\mathbb{R}^2$.

MSC2010: 37B20, 37B25.
Keywords: omega set, chain transitive homeomorphism, rigid space, Slovak space.

The dynamical systems we will consider are pairs (X, f), where $f : X \to X$ is a continuous map on a space X. If $x \in X$, then the sequence $x, f(x), f^2(x), \dots$ is the *trajectory* of x and $\omega f(x)$ is the set of limit points of the trajectory sequence, i.e., $y \in \omega f(x)$ if and only if $f^{n_k}(x) \to y$ for some sequence of integers $n_k \to \infty$.

The question when a system (X, f) can be embedded as a subsystem of some (Y, g) so that $X = \omega g(y)$ for some $y \in Y$ was answered by Dowker and Friedlander [1954] for homeomorphisms and by Sharkovsky [1965] in general. A system (X, f) is *f-connected* if, for any proper, nonempty, closed subset $U \subset X$, the intersection $f(U) \cap \overline{X \setminus U}$ is nonempty. They show that (X, f) can be embedded as the omega limit set in some larger system if and only if it is f-connected.

At the space level, the question now arises when a space X admits a map f so that (X, f) is an omega limit set subsystem. From the above results, this asks when X admits a map f with respect to which (X, f) is f-connected. Such a space X is called an *orbit enclosing omega limit set* when there exists a map f on X such that $X = \omega f(x)$ for some $x \in X$.

Considerable work, initiated by Sharkovsky, has been done on the related question of characterizing which subsets X of the unit interval I are omega limit sets for some map on the interval; see [Agronsky et al. 1989/90]. Notice that if X is a finite union of intervals or a Cantor set, then X is an orbit-enclosing omega limit set for some map f on X, so any extension via the Tietze extension theorem to a map on I will suffice. The result is more delicate for a general closed nowhere dense subset of the interval; see the elegant exposition in [Bruckner and Smítal 1992]. Later, Kolyada and Snoha [1992/93] extended these results by showing that a subset X of the unit interval is an omega limit set (or, equivalently, admits a chain transitive map) if and only if X is not a disjoint union of a finite number of nondegenerate intervals and a nonempty, countable set with the distance from this set to at least one of the intervals positive. Thus, this work is the first to consider the main question we will be addressing. For an early summary, see [Sharkovsky et al. 1989].

Recall that a *Peano space* is a compact, connected, locally connected space or, equivalently, a continuous image of the unit interval. In a pair of papers, Agronsky and Ceder [1991/92a; 1991/92b] proved that if X has finitely many components and each is a nontrivial, finite-dimensional Peano space, then X is an orbit enclosing omega limit set.

Our purpose here is to consider the related problem of when a space X admits a homeomorphism f so that (X, f) is the omega limit set in a larger system. As we will see, the results are somewhat different from the map case. First, we reinterpret the problem.

Given $\epsilon \geq 0$, a finite or infinite sequence $\{x_n \in X\}$ with at least two terms is an *ϵ-chain* for (X, f) if $d(f(x_k), x_{k+1}) \leq \epsilon$ for all terms x_k of the sequence (except the last one). The system (X, f) is called *chain transitive* when every pair of points of X can be connected by some finite ϵ-chain for every positive ϵ. A subset $A \subset X$ is called a *chain transitive subset* when it is closed and f-invariant (i.e., $f(A) = A$) and the subsystem (A, f) is chain transitive.

It is well known that any omega limit set is a chain transitive subset; see, e.g., [Akin 1993, Proposition 4.14]. On the other hand, as observed by Takens, it is easy to show that if (X, f) is chain transitive, then it can be embedded in a larger system in which it is an omega limit set [Akin 1993, Exercise 4.29]. We will review the proofs in Section 3. The construction uses a subset Y of $X \times [0, 1]$. In particular, if $X \subset [0, 1]^n$ then $Y \subset [0, 1]^{n+1}$. Applying the Tietze extension theorem in each coordinate, we can extend g to a continuous map on all of $[0, 1]^{n+1}$ and so obtain X as the omega limit set for a system on $[0, 1]^{n+1}$. Note, however, that even if g is a homeomorphism, we might not be able to extend it to a homeomorphism on $[0, 1]^{n+1}$.

These results are really just a restatement of the theorem of Dowker and Friedlander [1954].

To clarify the relationship between chain transitivity and f-connectedness, we recall the concept of an *attractor*, as described by Conley [1978] and with detailed exposition in [Akin 1993]. Call a closed set $U \subset X$ an *inward set* for f if $f(U)$ is contained in the interior U° or, equivalently, $f(U) \cap \overline{X \setminus U} = \varnothing$. Thus, (X, f) admits a proper, nonempty, inward subset if and only if it is not f-connected. If U is an inward set, then $A = \bigcap_{n=0}^{\infty} f^n(U)$ is called the associated attractor. For a number of equivalent descriptions of an attractor, see [Akin 1993, Theorem 3.3]. Theorem 4.12 of that paper says that (X, f) is chain transitive if and only if X is the only nonempty attractor and if and only if X is the only nonempty inward set. It follows that chain transitivity and f-connectedness are equivalent concepts.

The label "attractor" has been used for other ideas. Some authors refer to all omega limit sets as attractors. An attractor as defined above need not be chain transitive, and so there are attractors which are not omega limit sets. A more reasonable definition is that a set A is an attractor for f when it is f-invariant and, for every x in some neighborhood of A, we have $\omega f(x) \subset A$. This condition is necessary in order that A be an attractor à la Conley, but it is not sufficient. If X is the one-point compactification of $\mathbb{Z}$ and f is the extension to X of translation by 1 on $\mathbb{Z}$, then (X, f) is chain transitive. On the other hand, the point at infinity is the omega limit set of every point. By Theorem 3.6(a) of [Akin 1993], an f-invariant subset A is an attractor if and only if $\{x : \omega f(x) \subset A\}$ is a neighborhood of A and, in addition, A is stable, i.e., for every $\epsilon > 0$ there exists $\delta > 0$ such that if x is δ-close to A, then the forward orbit of x remains ϵ-close to A.

So the question we will address is when a space admits a chain transitive homeomorphism.

For a dynamical system (X, f), there exists $x \in X$ such that $\omega f(x) = X$ exactly when the system is topologically transitive with a recurrent transitive point. Thus, asking when X is an orbit-enclosing omega limit set is asking exactly when X admits f so that (X, f) is topologically transitive in this sense.

The identity map 1_X on X is chain transitive if and only if X is connected. Thus, a connected space admits a chain transitive homeomorphism and so is an omega limit set.

To illustrate the difference between the original problem and the homeomorphism version, consider the *tent map* on [0, 1] defined by

$$T(t) = \begin{cases} 2t & \text{for } 0 \le t \le \frac{1}{2}, \\ 1 - 2t & \text{for } \frac{1}{2} \le t \le 1, \end{cases} \tag{1-1}$$

which is well known to be topologically transitive. Now let $X_0 = [0, 1] \times \{0, 1\}$, and define f_0 on X_0 by

$$f_0(t, 0) = \langle t, 1\rangle \quad \text{and} \quad f_0(t, 1) = \langle T(t), 0\rangle, \tag{1-2}$$

so that $f_0^2 = T \times 1_{\{0,1\}}$.

Let (X, f) be the quotient system obtained by identifying the points $\langle 0, 1\rangle = \langle 1, 1\rangle$ in X. Thus, X is the disjoint union of a circle and an interval. It easily follows that X does not admit a chain transitive homeomorphism. On the other hand, f is a topologically transitive map.

Let (X_1, f_1) be the quotient system obtained from (X, f) by identifying the points $\langle 0, 0\rangle = \langle 1, 1\rangle$ in X_1. Thus, X_1 consists of a circle and an interval joined at a point. As X_1 is connected, the identity is a chain transitive homeomorphism. Since the points of the interval other than $\langle 1, 0\rangle$ separate the space and the points of the circle other than the intersection point with the interval do not, it easily follows that X_1 does not admit a topologically transitive homeomorphism. On the other hand, f_1 is a topologically transitive map.

If X contains a proper, clopen, nonempty, f-invariant set A, then we say that X is *f-decomposable*. In this case (X, f) is not chain transitive, for if ϵ is smaller than the distance from A to its complement, then any ϵ-chain which begins in A remains in A. With $H(X)$ the group of homeomorphisms on X, we say that X is *$H(X)$-decomposable* if there is a proper, clopen, nonempty subset A of X such that A is invariant for every homeomorphism on X. If X is not f-decomposable (or not $H(X)$-decomposable), we will call it *f-indecomposable* (resp. *$H(X)$-indecomposable*).

Clearly, if X is $H(X)$-decomposable, then it admits no chain transitive homeomorphism. For example, let $\mathrm{Iso}(X)$ denote the (possibly empty) set of isolated points of X. If the closure $\overline{\mathrm{Iso}(X)}$ is a proper, clopen, nonempty subset of X, then X is $H(X)$-decomposable and so admits no chain transitive homeomorphism. In the zero-dimensional case, this is the only obstruction. We prove a slightly more general result in Section 3.

Theorem 1.1. *If X is a space such that $\overline{\mathrm{Iso}(X)}$ is not a proper, clopen subset of X and such that the open set $X \setminus \overline{\mathrm{Iso}(X)}$ is empty or zero-dimensional, then X admits a chain transitive homeomorphism.*

Corollary 1.2. *If the isolated points are dense in X, then X admits a chain transitive homeomorphism.*

Thus, the problems which remain come from the nontrivial components. A clopen component is called an *isolated component*. Clearly, if the closure of the union of isolated components is a proper, clopen, nonempty subset of X, then X is $H(X)$-decomposable.

If $\mathcal{Y}$ is a set of connected spaces, let $C_{\mathcal{Y}}$ denote the closure of the union of those components of X which are homeomorphic to some element of $\mathcal{Y}$. If $C_{\mathcal{Y}}$ is a proper, clopen, nonempty subset, then X is $H(X)$-decomposable. For example, if X has finitely many components, then either they are all homeomorphic, in which case X admits a periodic chain transitive map, or X is $H(X)$-decomposable. Contrast this with the map result described above.

We say that X satisfies the *diameter condition on isolated components* if for every $\epsilon > 0$ there are only finitely many isolated components with diameter greater than ϵ.

Theorem 1.3. *If X satisfies the diameter condition and the union of the isolated components is dense in X, then either X is $H(X)$-decomposable or else X admits a chain transitive homeomorphism.*

The rest of Section 3 consists of counterexamples to reasonable conjectures. We construct:

- A space X that is $H(X)$-decomposable, with all components homeomorphic, and no isolated components.
- A space X that is $H(X)$-indecomposable but f-decomposable for every $f \in H(X)$. The space can be chosen with the isolated components all homeomorphic and with a dense union.
- A space X that is f-indecomposable for some $f \in H(X)$ but admits no chain transitive homeomorphism. The space can be chosen with the isolated components all homeomorphic and with a dense union.

These examples rule out the obvious extension of the above corollary. The isolated components can be dense and all homeomorphic to one another, but nonetheless the space admits no chain transitive homeomorphism.

Having considered when there are no chain transitive homeomorphisms, we consider in Section 4 this question: When is every homeomorphism on a space X chain transitive? For such a space X, the identity map 1_X is chain transitive, and so X must be connected.

In [de Groot and Wille 1958], rigid spaces were defined and Peano space examples were constructed. A space X is *rigid* if 1_X is the only homeomorphism on X, i.e., the homeomorphism group $H(X)$ is trivial. For a connected rigid space, it is trivially true that all homeomorphisms are chain transitive.

Using rigid spaces one can construct more interesting examples. Following [de Groot 1959], we can begin with a finitely generated group G and use rigid spaces instead of intervals as edges in the Cayley graph. If X is the one-point compactification of this fattened Cayley graph, then $H(X)$ is isomorphic to G and every homeomorphism is chain transitive. We obtain examples with nondiscrete homeomorphism group and even with the path components nontrivial.

In these cases, the homeomorphism group does not act in a topologically transitive manner on the space. Distinct points in each rigid piece are homeomorphically distinct. It is possible to obtain examples with all homeomorphisms chain transitive and with the homeomorphism group acting in a topologically transitive manner. These are built using the recent, beautiful construction in [Downarowicz et al. 2017] of Slovak spaces. A *Slovak space* X has $H(X)$ isomorphic to $\mathbb{Z}$ and every element in it other than the identity 1_X acts minimally on X. We call a space *Slovakian* if every homeomorphism other than 1_X is topologically transitive. Using such Slovakian spaces we construct a space X such that $H(X)$ is topologically transitive on X, every element of $H(X)$ is chain transitive, and the homeomorphism group of the Cantor set occurs as a closed, topological subgroup of $H(X)$.

2. Relation dynamics

It will be convenient to use the dynamics of closed relations, and so we briefly review the ideas from [Akin 1993]. Recall that our spaces X, Y, etc. are assumed to be nonempty, compact, metrizable spaces with a fixed metric chosen when necessary.

For spaces X, Y, a relation $R : X \to Y$ is a subset of $X \times Y$. The set R is a *relation on* X when $Y = X$. A map is a relation such that $R(x) = \{y : (x, y) \in R\}$ is a singleton set for every $x \in X$. Notice that we are following the set theory convention for which a map is the set sometimes referred to as the graph of the map. Thus, for example, the identity map 1_X is the diagonal set $\{(x, x) : x \in X\}$.

For $A \subset X$, the *image* $R(A)$ is defined to be $\bigcup_{x\in A} R(x)$. Equivalently, $R(A)$ is the projection to Y of $R \cap (A \times Y) \subset X \times Y$. The inverse $R^{-1} : Y \to X$ is defined to be $\{(y, x) : (x, y) \in R\}$. For $B \subset Y$, we let $R^*(B) = \{x \in X : R(x) \subset B\} = X \setminus R^{-1}(Y \setminus B)$. So $R^*(B) \subset R^{-1}(B) \cup R^*(\varnothing)$. If R is a map, then $R^*(B) = R^{-1}(B)$.

For example,

$$\overline{V}_\epsilon = \{(x, y) \in X \times X : d(x, y) \le \epsilon\}$$

is a relation on X with $\overline{V}_\epsilon(x)$ the closed ball of radius ϵ and center x. When $\epsilon = 0$, $\overline{V}_\epsilon = 1_X$, the identity map on X.

If $R : X \to Y$ and $S : Y \to C$ are relations, then the composition $S \circ R : X \to C$ is the image under the projection to $X \times C$ of the set $(R \times C) \cap (X \times S) \subset X \times Y \times C$. Thus, $(x, c) \in S \circ R$ if and only if there exists $y \in Y$ such that $(x, y) \in R$ and $(y, c) \in S$. Composition is associative, and $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.

A relation R on X is *reflexive* when $1_X \subset R$, *symmetric* when $R^{-1} = R$, and *transitive* when $R \circ R \subset R$.

For a relation R on X, we let $R^{n+1} = R^n \circ R$ and $R^{-n} = (R^{-1})^n$ for $n = 1, 2, \dots,$ and let R^0 be the identity 1_X. We define the *cyclic set* $|R| = \{x : (x, x) \in R\}$.

For a relation R on X a subset A of X is called *forward R-invariant* (or *R-invariant*) if $R(A) \subset A$ (resp. $R(A) = A$).

For a relation R on X, the *orbit relation* is $\mathcal{O}R = \bigcup_{n=1}^{\infty} R^n$, and the *orbit closure relation* $\mathcal{R}R$ is defined by $\mathcal{R}R(x) = \overline{\mathcal{O}R(x)}$ for all $x \in X$, so that $\mathcal{R}R = \{(x, y) : x \in X,\ y \in \overline{\mathcal{O}R(x)}\}$. The *wandering relation* is $\mathcal{N}R = \overline{\mathcal{O}R}$. Even when R is a continuous map, $\mathcal{R}R$ is usually not closed and so is a proper subset of $\mathcal{N}R$.

The *chain relation* is

$$\mathcal{C}R = \bigcap_{\epsilon > 0} \mathcal{O}(\overline{V}_\epsilon \circ R \circ \overline{V}_\epsilon). \tag{2-1}$$

Both $\mathcal{O}R$ and $\mathcal{C}R$ are transitive relations. Since $(R^n)^{-1} = (R^{-1})^n$, it follows that $\mathcal{O}(R^{-1}) = (\mathcal{O}R)^{-1}$, $\mathcal{N}(R^{-1}) = (\mathcal{N}R)^{-1}$ and $\mathcal{C}(R^{-1}) = (\mathcal{C}R)^{-1}$. These operators on relations are monotone, i.e., they preserve inclusions, and $\mathcal{O}$ and $\mathcal{C}$ are idempotent. That is,

$$\mathcal{O}(\mathcal{O}R) = \mathcal{O}R \quad \text{and} \quad \mathcal{C}(\mathcal{C}R) = \mathcal{C}R. \tag{2-2}$$

It then follows that

$$\begin{aligned} \mathcal{R}(\mathcal{O}R)(x) &= \overline{\mathcal{O}(\mathcal{O}R)(x)} = \overline{\mathcal{O}R(x)} = \mathcal{R}R(x), \\ \mathcal{N}(\mathcal{O}R) &= \overline{\mathcal{O}(\mathcal{O}R)} = \overline{\mathcal{O}R} = \mathcal{N}R, \\ \mathcal{C}R &\subset \mathcal{C}(\mathcal{O}R) \subset \mathcal{C}(\mathcal{C}R) = \mathcal{C}R. \end{aligned} \tag{2-3}$$

If R is a transitive relation on X, then $R \cap R^{-1}$ is a symmetric, transitive relation which restricts to an equivalence relation on $|R|$. We call the equivalence classes the *basic sets* of R.

A *closed relation* R is a closed subset of $X \times Y$. If R is a map, then it is continuous if and only if it is a closed relation, i.e., its graph is a closed set.

The composition of closed relations is closed, and the image of a closed set by a closed relation is closed. So if B is open in Y, then $R^*(B)$ is open in X. For any relation R, the extensions $\mathcal{N}R$ and $\mathcal{C}R$ are closed relations.

It is easy to see that $\mathcal{C}R = \mathcal{C}\overline{R}$, where $\overline{R}$ is the closure of R. If R is a closed relation, then $\mathcal{C}R = \bigcap_{\epsilon > 0} \mathcal{O}(\overline{V}_\epsilon \circ R)$; see [Akin 1993, Proposition 1.18]. If R is a closed relation on X, then the cyclic set $|R|$ is a closed subset of X. If R is a closed, transitive relation, the basic sets $\{R(x) \cap R^{-1}(x) : x \in |R|\}$ are closed.

For a sequence $\{A_n\}$ of closed sets, $\limsup\{A_n\} = \bigcap_n \overline{\bigcup_{k \geq n} A_k}$. We have the identity $\overline{\bigcup_n A_n} = \left(\bigcup_n A_n\right) \cup \limsup\{A_n\}$. If R is a closed relation on X and A is a closed subset of X, then we define $\omega R[A] = \limsup\{R^n(A)\}$.

If A is forward R-invariant and R is closed, then the sequence of closed sets $\{R^n(A)\}$ is decreasing and $\omega R[A]$ is the intersection. Furthermore, if $y \in \omega R[A]$, then $R^{-1}(y)$ meets every $\{R^n(A)\}$, and so by compactness it meets $\omega R[A]$ itself. It follows that $\omega R[A]$ is the maximum R-invariant subset of the closed, forward R-invariant set A.

If R is closed, then for $x \in X$ we let

$$\omega R(x) = \omega R[x] = \limsup\{R^n(x)\},$$

defining the *omega limit set relation* ωR on X. We also define $\Omega R = \limsup\{R^n\}$ for a general relation R, so ΩR is a closed relation. When R is closed, we have the identities

$$\mathcal{R}R = \mathcal{O}R \cup \omega R \quad \text{and} \quad \mathcal{N}R = \mathcal{O}R \cup \Omega R. \tag{2-4}$$

Since $\mathcal{C}R$ is closed and transitive, the sequence $\{(\mathcal{C}R)^n\}$ is decreasing with intersection $\Omega\mathcal{C}R = \omega\mathcal{C}R$.

If R is closed, then the following useful identities hold for the chain relation:

$$\begin{gathered} \mathcal{C}R = \mathcal{O}R \cup \Omega\mathcal{C}R, \\ R \cup ((\mathcal{C}R) \circ R) = \mathcal{C}R = R \cup (R \circ (\mathcal{C}R)); \end{gathered} \tag{2-5}$$

see [Akin 1993, Proposition 2.4(c), Proposition 1.11(d)].

It is not usually true that $(\omega R)^{-1} = \omega(R^{-1})$. We write αR for $\omega(R^{-1})$, defining the *alpha limit set relation*.

The points of $|\mathcal{O}R|$ are called *periodic points* for R, $|\mathcal{R}R|$ are the *recurrent points*, $|\mathcal{N}R|$ are the *nonwandering points*, and $|\mathcal{C}R|$ are the *chain recurrent points*. When R is a closed relation, we can apply the inclusion $|\mathcal{O}R| \subset |\omega R|$ to (2-4) and (2-5) to obtain $|\mathcal{R}R| = |\omega R|$, $|\mathcal{N}R| = |\Omega R|$ and $|\mathcal{C}R| = |\Omega\mathcal{C}R|$. We call x a *transitive point* for R when $\mathcal{R}R(x) = X$. We denote by $\mathrm{Trans}(R)$ the (possibly empty) set of transitive points.

On $|\mathcal{C}R|$, $\mathcal{C}R \cap \mathcal{C}R^{-1}$ is a closed equivalence relation. We call the equivalence classes, i.e., the basic sets for $\mathcal{C}R$, the *chain components* of R.

For a relation R on X, we will say that R is *minimal* when $\mathcal{R}R = X \times X$, *topologically transitive* when $\mathcal{N}R = X \times X$, and *chain transitive* when $\mathcal{C}R = X \times X$. We call R *central* when $|\mathcal{N}R| = X$ and *chain recurrent* when $|\mathcal{C}R| = X$. From (2-3) it follows that any one of these properties holds for R if and only if it holds for $\mathcal{O}R$.

Notice that if R is minimal, then X contains no proper closed, forward R-invariant subset. If R is a continuous map, then the converse is true as well since $\omega R(x)$ is then R-invariant and nonempty for every x. Thus, if R is a continuous map, it is minimal if and only if $X = \mathrm{Trans}(R)$.

A set U is *inward* for a closed relation R on X if U is closed and $R(U) \subset U^\circ$, the interior of U. Since $R(U)$ has a positive distance from the complement of U, it easily follows that U is inward for $\mathcal{C}R$ when it is inward for R. An inward set is therefore forward $\mathcal{C}R$-invariant. In general, a closed set A is forward $\mathcal{C}R$-invariant if and only if it has a neighborhood base consisting of inward sets; see [Akin 1993, Theorem 3.3(c)]. If U is an inward set, then $A_+ = \omega R[U] = \bigcap_{n\in\mathbb{N}} R^n(U)$ is the associated *attractor*. The set $X \setminus U^\circ$ is inward for R^{-1}, and $A_- = \omega(R^{-1})[X \setminus U^\circ]$ is the *dual repeller* for R. The pair A_+, A_- is called an *attractor-repeller pair*. If $x \in X \setminus (A_+ \cup A_-)$, then $\Omega\mathcal{C}R(x) \subset A_+$ and $\Omega\mathcal{C}R^{-1}(x) \subset A_-$; see [Akin 1993, Proposition 3.9].

Notice that a clopen set is inward if and only if it is forward invariant. If an inward set is invariant, then it is clopen.

Proposition 2.1. *Let R be a closed relation on X.*

(a) *If A_+, A_- is an attractor-repeller pair, then $|\mathcal{C}R| \subset A_+ \cup A_-$ and the intersections $A_+ \cap |\mathcal{C}R|$ and $A_- \cap |\mathcal{C}R|$ are each clopen in $|\mathcal{C}R|$.*

(b) *If $x \in X \setminus |\mathcal{C}R|$, then there exists an attractor-repeller pair A_+, A_- such that $x \notin A_+ \cup A_-$.*

(c) *If $x, y \in |\mathcal{C}R|$, then $y \in \mathcal{C}R(x)$ if and only if, for all attractors A, $x \in A$ implies $y \in A$.*

(d) *The space of chain components, i.e., the quotient space of $|\mathcal{C}R|$ by the equivalence relation $\mathcal{C}R \cap \mathcal{C}R^{-1}$, is a compact zero-dimensional metric space.*

Proof. (a) If $x \in X \setminus (A_+ \cup A_-)$, then $\Omega\mathcal{C}R(x) \subset A_+$, and so $x \notin \Omega\mathcal{C}R(x)$. If U is an inward set with $\omega R[U] = A_+$, then $U^\circ \cap |\mathcal{C}R| = A_+ \cap |\mathcal{C}R|$.

(b), (c) These are part of Proposition 3.11 of [Akin 1993].

(d) From (b) and (c) applied to R and R^{-1} it follows that every chain component C is the intersection of $\{A \cap |\mathcal{C}R| : A \text{ is an attractor or repeller and } C \subset A\}$. Hence, the attractors and repellers induce a clopen subbase on the space of chain components. □

For a relation R on a space X we say that X is *R-decomposable* if there is a proper, forward R-invariant decomposition $\{A_1, A_2\}$ of X, i.e., each is a nonempty, closed, forward R-invariant subset and $A_1 \cap A_2 = \emptyset$, $A_1 \cup A_2 = X$. Such a pair of proper, clopen sets is called an *R-decomposition*. Since the sets are clopen and forward R-invariant, they are inward for R. Hence, an R-decomposition is a $\mathcal{C}R$-decomposition. So the notion of decomposability is the same for R, $\mathcal{O}R$, $\mathcal{R}R$, $\mathcal{N}R$ and $\mathcal{C}R$. If no such decomposition exists, X is said to be *R-indecomposable*.

For any relation R on X let $R_\pm = R \cup 1_X \cup R^{-1}$. This is a reflexive and symmetric relation on X so that $\mathcal{O}(R_\pm)$ and $\mathcal{C}(R_\pm)$ are equivalence relations on X.

Proposition 2.2. *For a relation R on X, the following conditions are equivalent:*

(i) *The space X is R-indecomposable.*

(ii) *The space X is $R_\pm$-indecomposable.*

(iii) *The relation $R_\pm$ is chain transitive.*

Proof. (i) $\Leftrightarrow$ (ii): If a set is forward R-invariant, then its complement is forward R^{-1}-invariant. Hence, an R-decomposition is an $R_\pm$-decomposition. Clearly, the reverse is true since $R \subset R_\pm$.

(iii) $\Rightarrow$ (ii): If $\{A_1, A_2\}$ is a decomposition, then each is inward for $R_\pm$ and therefore forward $\mathcal{C}R_\pm$-invariant. Hence, $R_\pm$ is not chain transitive.

(ii) $\Rightarrow$ (iii): If $R_\pm$ is not chain transitive, the equivalence relation $\mathcal{C}R_\pm$ on X has more than one equivalence class. By Proposition 2.1(d) the space of equivalence classes is zero-dimensional. Hence, there is a pair of disjoint, nonempty, clopen sets $\{A_1, A_2\}$ which cover X such that each is a union of equivalence classes. Since each equivalence class is $\mathcal{C}R_\pm$-invariant, it follows that $\{A_1, A_2\}$ is a $R_\pm$-decomposition. $\square$

Definition 2.3. Let $\pi : X_1 \to X_2$ be a continuous map between spaces, and let R_j be a relation on X_j for $j = 1, 2$. We say that π *maps R_1 to R_2* if $\pi \circ R_1 \subset R_2 \circ \pi$ and that π *is a semiconjugacy from R_1 to R_2* if $\pi \circ R_1 = R_2 \circ \pi$.

Proposition 2.4. *Let $\pi : X_1 \to X_2$ be a continuous map between spaces, and let R_j be a relation on X_j for $j = 1, 2$.*

(a) *The function π maps R_1 to R_2 if and only if $(\pi \times \pi)(R_1) \subset R_2$. If π is a semiconjugacy from R_1 to R_2 and $\pi(X_1) = X_2$, i.e., π is surjective, then $(\pi \times \pi)(R_1) = R_2$.*

(b) *If R_j is a mapping for $j = 1, 2$, then π is a semiconjugacy from R_1 to R_2 if it maps R_1 to R_2.*

(c) *If π maps R_1 to R_2, then π maps R_1^n to R_2^n for all $n \in \mathbb{Z}$ and maps $\mathcal{A}R_1$ to $\mathcal{A}R_2$ for $\mathcal{A} = \mathcal{O}, \mathcal{R}, \mathcal{N}$ and $\mathcal{C}$.*

(d) *Assume π is surjective and maps R_1 to R_2. If R_1 is minimal, topologically transitive, central, chain transitive or chain recurrent, then R_2 satisfies the corresponding property.*

(e) *If π is a semiconjugacy from R_1 to R_2 and π is an open map, then π is a semiconjugacy from $\mathcal{A}R_1$ to $\mathcal{A}R_2$ for $\mathcal{A} = \mathcal{O}, \mathcal{R}, \mathcal{N}$ and $\mathcal{C}$.*

Proof. (a) Firstly, π maps the relation R_1 to R_2 if and only if $(x, y) \in R_1$ implies $(\pi(x), \pi(y)) \in R_2$, and so the first equivalence is clear. The second result is easy to check.

(b) An inclusion between functions is an equation.

(c) Given $\epsilon > 0$, there exists $\delta > 0$ such that $(\pi \times \pi)(\overline{V}_\delta) \subset V_\epsilon$ since π is uniformly continuous. If $\{(x_n, y_n) \in R_1\}$ is a finite or infinite sequence with $(y_n, x_{n+1}) \in \overline{V}_\delta$, then $\{(\pi(x_n), \pi(y_n)) \in R_2\}$ with $(\pi(y_n), \pi(x_{n+1})) \in \overline{V}_\epsilon$. That is, δ-chains for R_1 are mapped to ϵ-chains for R_2. It follows that π maps $\mathcal{C}R_1$ to $\mathcal{C}R_2$. The others are easy to check.

(d) If π is surjective and maps $\mathcal{A}R_1$ to $\mathcal{A}R_2$, then $X_1 \times X_1 = \mathcal{A}R_1$ implies $X_2 \times X_2 = \mathcal{A}R_2$, and $1_{X_1} \subset \mathcal{A}R_1$ implies $1_{X_2} \subset \mathcal{A}R_2$.

(e) The function $1_{X_1} \times \pi$ maps 1_{X_1} to $\pi \subset X_1 \times X_2$. If π is open and $\epsilon > 0$, then $1_{X_1} \times \pi$ maps $\overline{V}_\epsilon$ to a neighborhood of π and so contains $\overline{V}_\delta \circ \pi$ for some $\delta > 0$. That is, $\overline{V}_\delta \circ \pi \subset \pi \circ \overline{V}_\epsilon$. If $\{(u_n, v_n) \in R_2\}$ is a finite or infinite sequence with $(v_n, u_{n+1}) \in \overline{V}_\delta$ and $\pi(x_n) = u_n$, then because π is a semiconjugacy on R_1, there exists $y_n \in X_1$ such that $(x_n, y_n) \in R_1$ and $\pi(y_n) = v_n$, and there exists x_{n+1} such that $(y_n, x_{n+1}) \in \overline{V}_\epsilon$ and $\pi(x_{n+1}) = u_{n+1}$. Thus, every δ-chain for R_2 can be lifted to an ϵ-chain for R_1 with a given initial lift. That is, π is a semiconjugacy from $\mathcal{C}R_1$ to $\mathcal{C}R_2$. The cases of $\mathcal{A} = \mathcal{O}, \mathcal{R}$ and $\mathcal{N}$ are similar. In fact, for $\mathcal{O}$ and $\mathcal{R}$ it is not necessary that the map be open. $\square$

Remark. Suppose $R_j = 1_{X_j}$ with $X_2 = [0, 1]$ and $X_1 = \{-1\} \cup [0, 1]$, and let $\pi : X_1 \to X_2$ be an extension of the identity on $[0, 1]$, mapping -1 to some point of $[0, 1]$. The map π is a semiconjugacy from R_1 to R_2 and maps $\mathcal{C}R_1$ onto $\mathcal{C}R_2$, but it is not a semiconjugacy from $\mathcal{C}R_1$ to $\mathcal{C}R_2$.

As indicated in the Introduction, our concern is with homeomorphisms. However, we will apply this machinery to three different relations.

Let $H(X)$ be the homeomorphism group of X. For $f \in H(X)$, we define $f_\pm$ as the closed relation $f \cup 1_X \cup f^{-1}$. Clearly, we have

$$\begin{aligned}
\mathcal{O}f &= \{f^n : n \in \mathbb{N}\},\\
\mathcal{O}f_\pm &= \{f^n : n \in \mathbb{Z}\} = \mathcal{O}f \cup 1_X \cup \mathcal{O}f^{-1},\\
\mathcal{R}f(x) &= \overline{\{f^n(x) : n \in \mathbb{N}\}},\\
\mathcal{R}f_\pm(x) &= \overline{\{f^n(x) : n \in \mathbb{Z}\}} \quad \text{for } x \in X,\\
\mathcal{N}f_\pm &= \mathcal{N}f \cup 1_X \cup \mathcal{N}f^{-1},\\
\mathcal{R}f_\pm &= \mathcal{R}f \cup 1_X \cup \mathcal{R}(f^{-1}).
\end{aligned} \tag{2-6}$$

On the other hand, $\mathcal{C}f_\pm$ is in general larger than $\mathcal{C}f \cup 1_X \cup \mathcal{C}f^{-1}$. The latter relation need not be transitive. See Example 2.8 below.

Since $1_X \subset \mathcal{O}f_\pm$, every point is recurrent for $f_\pm$.

An action of a group G on X is a homomorphism $\rho : G \to H(X)$. If G is a subgroup of $H(X)$, then G acts on X by evaluation. That is, ρ is the inclusion.

With the action understood we let $h_G = \bigcup\{f \in \rho(G)\}$. This is just the orbit relation of the action of G on X. It is an equivalence relation but usually not closed. Since it is an equivalence relation, $\mathcal{O}h_G = h_G$ and $\mathcal{N}h_G = \overline{h_G}$. Since h_G is reflexive, every point is recurrent for h_G.

If G is the cyclic group generated by a homeomorphism f, then $h_G = \mathcal{O} f_{\pm}$.

When $G = H(X)$, we write h_X for h_G. Clearly, $H(X)$-decomposability as described in the Introduction is just h_X-decomposability.

For a space X, let $\mathrm{Iso}(X)$ denote the set of isolated points of X.

Proposition 2.5. *Let $f \in H(X)$ and $\rho : G \to H(X)$ be an action of a group G on X. Let $\mathcal{B}$ be a countable base of nonempty open sets for X, and let $\mathcal{B}^*$ be the collection of finite covers of X by elements of $\mathcal{B}$.*

(a) *The homeomorphism f is central if and only if the G_δ set of recurrent points, $|\omega f|$, is dense:*

$$|\omega f| = \bigcap_{\mathcal{A} \in \mathcal{B}^*,\, n \in \mathbb{N}} \ \bigcup_{U \in \mathcal{A},\, i \geq n} \{U \cap f^{-i}(U)\}. \tag{2-7}$$

If f is central, then any isolated point x of X is a periodic point for f.

(b) *The homeomorphism f is topologically transitive if and only if the set of transitive points,* $\mathrm{Trans}(f)$, *is nonempty, in which case*

$$\mathrm{Trans}(f) = \{x : \omega f(x) = X\} \tag{2-8}$$

$$= \bigcap_{U \in \mathcal{B},\, n \in \mathbb{N}} \ \bigcup_{i \geq n} \{f^{-i}(U)\} \tag{2-9}$$

is a dense G_δ subset of X. If f is topologically transitive, then f^{-1} is topologically transitive. If f is topologically transitive, then X is perfect or consists of a single periodic orbit for f.

(c) *The relation $f_{\pm}$ is topologically transitive if and only if the set of transitive points,* $\mathrm{Trans}(f_{\pm})$, *is nonempty, in which case*

$$\mathrm{Trans}(f_{\pm}) = \bigcap_{U \in \mathcal{B}} \ \bigcup_{n \in \mathbb{Z}} \{f^{-n}(U)\} \tag{2-10}$$

is a dense G_δ subset of X. If $f_{\pm}$ is topologically transitive and x is an isolated point of X, then

$$\mathrm{Trans}(f_{\pm}) = \mathcal{O} f_{\pm}(x) = \mathrm{Iso}(X). \tag{2-11}$$

Thus, if $f_{\pm}$ is topologically transitive, then either $\mathrm{Iso}(X) = \varnothing$, *so X is perfect, or else* $\mathrm{Iso}(X)$ *is dense. If* $\mathrm{Iso}(X)$ *is finite and nonempty, then X consists of a single periodic orbit.*

(d) *The following are equivalent*:

(i) f *is topologically transitive*;

(ii) $f_\pm$ *is topologically transitive and* f *is central*;

(iii) $f_\pm$ *is topologically transitive and* X *is either perfect or consists of a single periodic orbit for* f.

In that case, $\mathrm{Trans}(f_\pm) = \mathrm{Trans}(f) \cup \mathrm{Trans}(f^{-1})$.

(e) *The relation* h_G *is topologically transitive if and only if the set of transitive points*, $\mathrm{Trans}(h_G)$, *is nonempty*, *in which case*

$$\mathrm{Trans}(h_G) = \bigcap_{U \in \mathcal{B}} \bigcup_{f \in G} \{f^{-1}(U)\} \tag{2-12}$$

is a dense G_δ *subset of* X.

Proof. For a relation R on X and subsets $U, V \subset X$, let

$$\begin{aligned} N_R(U, V) &= \{n \in \mathbb{N} : R^n(U) \cap V \neq \varnothing\} \\ &= \{n \in \mathbb{N} : U \cap R^{-n}(V) \neq \varnothing\}. \end{aligned} \tag{2-13}$$

Clearly, R is central if and only if $N_R(U, U) \neq \varnothing$ for all nonempty open subsets U, and R is topologically transitive if and only if $N_R(U, V) \neq \varnothing$ for all nonempty open subsets U, V. If $R = f \in H(X)$ and $N_f(U, V)$ is finite for some open U, V, then with $n = \max N_f(U, V)$ we let $W = U \cap f^{-n}(V)$. Then W is a nonempty open set and $N_f(W, V) = \varnothing$. Hence, if f is central, then $N_f(U, U)$ is infinite, and if f is topologically transitive, then $N_f(U, V)$ is infinite for all nonempty open U, V.

The equations (2-7)–(2-13) are easy to check, and density follows from the Baire category theorem.

Now assume that $f_\pm$ is topologically transitive and that x is an isolated point of X. If V is a nonempty open set, then $N_{f_\pm}(\{x\}, V) = \varnothing$ unless V meets the orbit $\mathcal{O} f_\pm(x)$. Hence, $x \in \mathrm{Trans}(f_\pm)$. If y is another isolated point, then $N_{f_\pm}(\{x\}, \{y\}) \neq \varnothing$ implies that y is in the $f_\pm$ orbit of x. Thus, (2-11) holds. In particular, $\mathrm{Iso}(X)$ is dense if it is nonempty. If it is finite and nonempty, then $X = \mathrm{Iso}(X)$ because the latter is closed and dense. Since the elements of $\mathrm{Iso}(X)$ lie on a single orbit, X consists of a single periodic orbit.

If f is central and $x \in \mathrm{Iso}(X)$, then $N_f(\{x\}, \{x\}) \neq \varnothing$ if and only if x is a periodic point. Thus, if $f_\pm$ is transitive and f is central, then either X is perfect or it consists of a single periodic orbit, proving the implication (ii) $\Rightarrow$ (iii) in (d). Since (i) obviously implies (ii) in (d), it follows that if f is topologically transitive, then either X is perfect or it consists of a single periodic orbit. So if $\mathcal{R} f(x) = X$, then $\overline{\{f^k(x) : k \geq n\}} = X$ for every $n \in \mathbb{N}$. Intersecting, we see that $\omega f(x) = X$. Thus, $\mathrm{Trans}(f) = \{x : \omega f(x) = X\}$.

Finally, if X is a single periodic orbit, then f is topologically transitive and every point is a transitive point for f and f^{-1}. Now assume that X is perfect and that $f_{\pm}$ is transitive. If $x \in \mathrm{Trans}(f_{\pm})$, then $\overline{\{f^k(x) : k \in \mathbb{Z},\ |k| \geq n\}} = X$ for all $n \in \mathbb{N}$. Intersecting, we obtain that $\omega f(x) \cup \alpha f(x) = X$. In particular, x is in one of these. If x is contained in a closed, invariant set A like $\omega f(x)$ or $\alpha f(x)$, then $X = \mathcal{R}(f_{\pm})(x) \subset A$. Thus, $x \in \mathrm{Trans}(f) \cup \mathrm{Trans}(f^{-1})$, so either f or f^{-1} is topologically transitive. But $\mathcal{N}(f^{-1}) = (\mathcal{N}f)^{-1}$ implies that f^{-1} is topologically transitive if f is. □

Remark. There are various, slightly conflicting, definitions of topological transitivity. These are sorted out in [Akin and Carlson 2012]. We are following [Akin 1993].

Lemma 2.6. (a) *Let $f \in H(X)$ and $x \in X$. The sets $\omega f(x)$ and $\alpha f(x)$ are f-invariant. If $x \in |\mathcal{C}f|$, then $\mathcal{C}f(x)$ and $\mathcal{C}f^{-1}(x)$ are f-invariant.*

(b) *The chain components of 1_X are the components of X.*

Proof. (a) For a bijection f on X, a subset A is f-invariant if and only if it is f^{-1}-invariant if and only if it is forward invariant for both f and f^{-1}.

When f is a continuous map on X, then $y \in \omega f(x)$ when there is a subsequence $\{f^{n_i}(x)\}$ of the orbit sequence which converges to y. Then $\{f^{n_i+1}(x)\}$ converges to $f(y)$, and if a subsequence of $\{f^{n_i-1}(x)\}$ converges to z, then $f(z) = y$. That is, $\omega f(x)$ is a nonempty, closed, f-invariant subset of X when f is a continuous map on X. So when f is a homeomorphism, the same is true for $\alpha f(x)$.

For $f \in H(X)$ we obtain from the identity (2-5) that $1_X \cup \mathcal{C}f = f^{-1} \circ \mathcal{C}f$. If $x \in \mathcal{C}f(x)$, then $\mathcal{C}f(x) = \{x\} \cup \mathcal{C}f(x) = f^{-1}(\mathcal{C}f(x))$. That is, $\mathcal{C}f(x)$ is f^{-1}-invariant and so is f-invariant. Since $\mathcal{C}(f^{-1}) = (\mathcal{C}f)^{-1}$ the same is true for $\mathcal{C}f^{-1}(x)$.

(b) The space of chain components being zero-dimensional by Proposition 2.1(d), every connected set contained in $|\mathcal{C}R|$ is entirely contained in a single chain component. In particular, when R is 1_X, every component is a subset of a chain component. On the other hand, when $R = 1_X$, every clopen subset is inward for R and so contains any chain component that it meets. It follows that the components are the chain components. □

Proposition 2.7. *Let $f \in H(X)$ and $\rho : G \to H(X)$ be an action of a group G on X.*

(a) *The following are equivalent*:
   (i) *X is f-indecomposable*;
   (ii) *X is $f_{\pm}$-indecomposable*;
   (iii) *$f_{\pm}$ is chain transitive.*

(b) *The relation h_G is chain transitive if and only if h_G is indecomposable.*

(c) *If f is chain recurrent, then $\mathcal{C}f = \mathcal{C}(f_{\pm})$.*

(d) *The following are equivalent*:
  (i) f *is chain transitive*;
  (ii) f *is chain recurrent and* $f_{\pm}$ *is chain transitive*;
  (iii) f *is chain recurrent and indecomposable.*

(e) *If* X *is connected, then* f *is chain transitive if and only if it is chain recurrent. In particular,* 1_X *is chain transitive.*

Proof. (a), (b) Since $h_G = (h_G)_{\pm}$, these results are a special case of Proposition 2.2 applied with $R = f$ and $R = h_G$.

(c) If f is chain recurrent, then $|\mathcal{C}f| = X$ and $1_X \subset \mathcal{C}f$. For any $x \in X$, we have $x \in \mathcal{C}f(x)$, and this set is f-invariant by Lemma 2.6. Hence, $f^{-1}(x) \in \mathcal{C}f(x)$. Thus, $f^{-1} \subset \mathcal{C}f$. Since $f \subset f_{\pm} \subset \mathcal{C}f$, it follows from (2-2) that $\mathcal{C}f \subset \mathcal{C}(f_{\pm}) \subset \mathcal{C}\mathcal{C}f = \mathcal{C}f$.

(d) If f is chain transitive, then it is clearly chain recurrent, and $f \subset f_{\pm}$ implies that $f_{\pm}$ is chain transitive. The converse follows from (c). This proves the equivalence of (i) and (ii). The equivalence of (ii) and (iii) follows from (a).

(e) If X is connected, then by Lemma 2.6(b) X consists of a single chain component for 1_X, and so 1_X is chain transitive. So if f is chain recurrent, then $X \times X = \mathcal{C}1_X \subset \mathcal{C}\mathcal{C}f = \mathcal{C}f$ by (2-2) again. □

Example 2.8. On $\mathbb{Z}$ let t be the translation bijection given by $n \mapsto n+1$. Let $\mathbb{Z}^*$ be the one-point compactification adjoining the point $\pm\infty$ to $\mathbb{Z}$, and let $\mathbb{Z}^{**}$ be the two-point compactification adjoining the points $-\infty, +\infty$ to $\mathbb{Z}$. Let t^* and t^{**} be the homeomorphisms extending t to $\mathbb{Z}^*$ and $\mathbb{Z}^{**}$, respectively. Both $t^*_{\pm}$ and $t^{**}_{\pm}$ are topologically transitive with $\mathbb{Z}$ the orbit of isolated points, and so $\mathbb{Z} = \mathrm{Trans}(t^*_{\pm}) = \mathrm{Trans}(t^{**}_{\pm})$. Of course, both $t^*_{\pm}$ and $t^{**}_{\pm}$ are chain transitive, but t^* is also chain transitive while t^{**} is not.

Let X be the quotient space of $Z^{**} \times \{0, 1\}$ with the fixed points $\langle +\infty, 0\rangle$ and $\langle +\infty, 1\rangle$ identified. Let f be the homeomorphism on X induced by $t^{**} \times 1_{\{0,1\}}$. Clearly, $f_{\pm}$ is chain transitive. But $\mathcal{C}f \cup 1_X \cup \mathcal{C}f^{-1}$ is contained in the image of $(Z^{**} \times \{0\})^2 \cup (Z^{**} \times \{1\})^2$ in X^2 and so is a proper subset of $\mathcal{C}(f_{\pm}) = X^2$. Furthermore, it is easy to check that in this case $\mathcal{C}(f \cup 1_X) = (\mathcal{C}f) \cup 1_X$. Hence, $R = f \cup 1_X$ is a closed relation such that R is chain recurrent and $R_{\pm}$ is chain transitive, but R is not chain transitive. Thus, (c) and (d) in Proposition 2.7 do not extend to general relations. □

Finally, recall the *uniqueness of Cantor*: any compact, perfect, zero-dimensional, metrizable space is homeomorphic to the Cantor set in [0, 1]. We will call any such space a Cantor set. We will need the following well-known result, and we provide a brief sketch of the proof.

Proposition 2.9. *For any space* X *there exists a surjective continuous map* $\pi : C \to X$ *with* C *a Cantor set.*

Proof. With $\mathcal{B}$ a countable basis for X, let Z be the closure in $X \times \{0, 1\}^{\mathcal{B}}$ of the set of pairs $\{(x, z) : z_U = 1 \Leftrightarrow x \in U\}$. The projection to X is clearly onto, and because $\mathcal{B}$ is a basis, the projection to $\{0, 1\}^{\mathcal{B}}$ is easily seen to be injective. It follows that Z is compact and zero-dimensional. If C_0 is a Cantor set, then $C = Z \times C_0$ is perfect as well as zero-dimensional and so is a Cantor set. Let π be the composition of projections $C \to Z \to X$. $\square$

3. Spaces which admit chain transitive maps

We begin with the relationship between omega limit sets and chain transitive subsets which was described in the Introduction.

Proposition 3.1. *If f is a homeomorphism on a space X and $x \in X$, then $\omega f(x)$ and $\alpha f(x)$ are chain transitive subsets, i.e., the restriction of f to each of these nonempty, closed, invariant sets is chain transitive.*

Proof. Let $y, y' \in \omega f(x)$ and let $\epsilon > 0$. Choose $\delta > 0$ an $\epsilon/2$-modulus of uniform continuity for f with $\delta < \epsilon/2$. There exists $N \in \mathbb{N}$ such that $n \geq N$ implies $f^n(x) \in \overline{V}_\delta(\omega f(x))$. There exists $n \geq N$ such that $d(f^n(x), y) < \delta$ and $k \in \mathbb{N}$ such that $d(f^{n+k}(x), y') < \delta$. For $j = 0, \dots, k$ choose $y_j \in \omega f(x)$ such that $d(f^{n+j}(x), y_j) \leq \delta$ with $y_0 = y, y_k = y'$. Hence, $d(f^{n+j+1}(x), f(y_j)) \leq \epsilon/2$ and so $d(y_{j+1}, f(y_j)) \leq \epsilon$. That is, $\{y_j\}$ is an ϵ-chain from y to y'. It follows that $\omega f(x)$ is a chain transitive subset. Applying the result to f^{-1}, we see that $\alpha f(x) = \omega(f^{-1})(x)$ is a chain transitive subset as well, since f is chain transitive if and only if f^{-1} is. $\square$

We prove, conversely, in Theorem 3.13(a) below that a chain transitive homeomorphism is the restriction to an omega limit set of a homeomorphism in a larger system.

In the constructions which follow, we will repeatedly use the process of *attachment*. Assume that A is a nonempty, closed, nowhere dense subset of a space X and that $h : A \to B$ is a continuous surjection. We may assume that X and B are disjoint. Otherwise, replace X by $X \times \{0\}$ and B by $B \times \{1\}$. Define X/h, the space with X *attached to B via h* as follows: Let $\tilde{h}$ denote the continuous retraction $h \cup 1_B : A \cup B \to B$. Let $E_h = 1_X \cup (\tilde{h}^{-1} \circ \tilde{h}) = 1_X \cup (\tilde{h} \times \tilde{h})^{-1}(1_B)$, a closed equivalence relation on $X \cup B$. Let X/h be the quotient space with projection $q_h : X \cup B \to X/h$. Since q_h is injective on B, we may regard it as an identification and so regard B as a subset of X/h. Furthermore, q_h restricts to a surjection $X \to X/h$ which maps A onto B via h and which is a homeomorphism between the dense open sets $X \setminus A \subset X$ and $(X/h) \setminus B \subset X/h$.

When B is a singleton, we write $q_A : X \to X/A$ for the quotient map and describe the result as *smashing A to a point*.

We will use some results of E. R. Lorch [1981; 1982] (see also [Tsankov 2006]), which we will briefly review.

For a locally compact space W, a *compactification* of W is a compact space Y together with a dense embedding of W into Y, i.e., a homeomorphism of W onto a dense subset of Y. Because W is locally compact, its image is an open, dense subset of Y, so $X = Y \setminus W$ is a nowhere dense, closed subset of Y. Reversing the point of view, we call Y an *extension of* X if Y is a compact space and X is a closed, nowhere dense subset of Y.

By a *pair of spaces* (Y, X) we will mean a space and a closed subset, respectively. Recall our default assumption that a space is a nonempty, compact, metrizable space. We will call (Y, X) an *extension pair* when Y is an extension of X, i.e., when X is nowhere dense in Y.

A continuous map $f : Y_1 \to Y_2$ is a map of pairs $f : (Y_1, X_1) \to (Y_2, X_2)$ when $f(X_1) \subset X_2$. If $X_1 = X_2 = X$, we will say that $f : (Y_1, X) \to (Y_2, X)$ is a map *rel* X if it restricts to the identity on X and if, in addition, $f(Y_1 \setminus X) \subset Y_2 \setminus X$. These conditions imply that $1_X = (f \times f)^{-1}(1_X)$. So by intersecting over $\delta > 0$ and using compactness, we see that for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$\begin{gathered}\text{if } (u, v) \in Y_1 \times Y_1 \text{ and } d_2(f(u), f(v)),\, d_2(f(u), X),\, d_2(f(v), X) \le \delta,\\ \text{then } (u, v) \in \overline{V}^{d_1}_{\epsilon},\\ \text{and so } f^{-1}(\overline{V}^{d_2}_{\delta}(x)) \subset \overline{V}^{d_1}_{\epsilon}(x) \text{ for all } x \in X,\end{gathered} \tag{3-1}$$

where d_1, d_2 are metrics on Y_1 and Y_2. Note that in considering different extensions Y_1, Y_2 of the same space X we do not assume that the metrics d_1 and d_2 agree on X, although they are, of course, uniformly equivalent on X.

Definition 3.2. We call $X^{(p)}$ an *isolated point extension* of X, or just a *point extension* of X, if $X^{(p)}$ is an extension of X with each point of $X^{(p)} \setminus X$ isolated. We then call $(X^{(p)}, X)$ a *point extension pair*.

A pair (Y, A) is a point extension pair if and only if Y is infinite, $\mathrm{Iso}(Y)$ is dense and $A = Y \setminus \mathrm{Iso}(Y)$. Since Y is separable, $\mathrm{Iso}(Y)$ is denumerable, and so Lorch uses the term *denumerable extension* instead of point extension. Thus, Y is a compactification of the denumerable discrete set $\mathrm{Iso}(Y)$.

Lemma 3.3. *Let $\pi : (Y, X) \to (X^{(p)}, X)$ be a surjective map of pairs rel X with $X^{(p)}$ a point extension of X. For every $\epsilon > 0$, the set $\{x \in X^{(p)} : \operatorname{diam} \pi^{-1}(x) > \epsilon\}$ is finite.*

Proof. By (3-1) there exists $\delta > 0$ such that for $x \in X^{(p)}$ with $x \in \overline{V}_{\delta}(X)$, $\operatorname{diam} \pi^{-1}(x) \le \epsilon$. So the set $\{x \in X^{(p)} : \operatorname{diam} \pi^{-1}(x) > \epsilon\}$ is contained in the complement of the open δ-neighborhood of X in $X^{(p)}$. This is a compact set of isolated points and so is finite. □

For a point extension $(X^{(p)}, X)$, we define a *canonical retraction* $r : X^{(p)} \to X$ so that for all $x \in X^{(p)}$

$$d(x, r(x)) = d(x, X). \tag{3-2}$$

The choice depends on the metric. Even for a fixed metric there may be more than one closest to X point $r(x)$. For each x we fix a choice to define r. Clearly,

$$d(r(x_1), r(x_2)) \le d(x_1, X) + d(x_1, x_2) + d(x_2, X). \tag{3-3}$$

For every $\epsilon > 0$ there are only finitely many points $x \in X^{(p)}$ with $d(x, X) \ge \epsilon$, so continuity of r at the points of X follows. Continuity at the isolated points is trivial.

Notice that if N_0 is a cofinite subset of $X^{(p)} \setminus X$, then the closure of N_0 contains X, so $r(N_0)$ is dense in X.

We recall the elegant proof of *Lorch's uniqueness theorem* [1981, Proposition 10].

Theorem 3.4. *Every space X has an essentially unique isolated point extension. That is, if (Y, X) and (Y', X) are point extension pairs, then there is a homeomorphism $f : (Y, X) \to (Y', X)$ rel X.*

Proof. Let $\{x_n : n \in \mathbb{N}\}$ be a sequence of not necessarily distinct points in X such that $\{x_n : n \ge N\}$ is dense in X for all $N \in \mathbb{N}$. Let

$$X^{(p)} = X \times \{0\} \cup \{(x_n, n^{-1}) : n \in \mathbb{N}\} \subset X \times [0, 1],$$

and identify X with $X \times \{0\}$. Clearly, $(X^{(p)}, X)$ is a point extension pair, so X has at least one point extension.

Let (Y, X) and (Y', X) be point extension pairs. Fix metrics d and d' on Y and Y', respectively. Define a metric d'' on X as the pointwise maximum of d and d'. The three metrics d, d' and d'' are uniformly equivalent on X. Let $r : Y \to X$ and $r' : Y' \to X$ be canonical retractions. Let $N = \mathrm{Iso}(Y) = Y \setminus X$ and $N' = \mathrm{Iso}(Y') = Y' \setminus X$. Use a counting of N and N' to impose orderings which are order isomorphic to $\mathbb{N}$ and so are well-orderings.

Let a_1 be the first element of N, and choose b_1 to be the first element of N' which satisfies

$$d''(r(a_1), r'(b_1)) < d(a_1, X). \tag{3-4}$$

Let b_2 be the first element of $N' \setminus \{b_1\}$, and choose a_2 to be the first element of $N \setminus \{a_1\}$ such that

$$d''(r(a_2), r'(b_2)) < d'(b_2, X). \tag{3-5}$$

Proceed inductively. If n is even, let a_{n+1} be the first element of $N \setminus \{a_1, \dots, a_n\}$, and if n is odd, let b_{n+1} be the first element of $N' \setminus \{b_1, \dots, b_n\}$. We can then choose b_{n+1} or a_{n+1} so that

$$d''(r(a_{n+1}), r'(b_{n+1})) < \max\{d(a_{n+1}, X), d'(b_{n+1}, X)\}. \tag{3-6}$$

By construction, $\{a_n\}$ and $\{b_n\}$ are renumberings of the sets N and N'. Define the mapping $f : Y \to Y'$ as an extension of the identity on X by putting $f(a_n) = b_n$ for all n. Let $m_n = \max\{d(a_n, X), d'(b_n, X)\}$. Observe that $m_n \to 0$ as $n \to \infty$.

Continuity of f is clear at the isolated points. Suppose $x \in X$ and $a_{n_i} \to x$ so that $n_i \to \infty$. Then

$$\begin{aligned} d(r(a_{n_i}), x) &\le d(a_{n_i}, x) + d(r(a_{n_i}), a_{n_i}) \\ &\le d(a_{n_i}, x) + m_{n_i} \to 0, \end{aligned} \tag{3-7}$$

and so $d'(r(a_{n_i}), x) \to 0$. Hence,

$$\begin{aligned} d'(b_{n_i}, x) &\le d'(r(b_{n_i}), b_{n_i}) + d'(r(b_{n_i}), r(a_{n_i})) + d'(r(a_{n_i}), x) \\ &\le 2m_{n_i} + d'(r(a_{n_i}), x) \to 0. \end{aligned} \tag{3-8}$$

Continuity of f follows. The result for f^{-1} is similar and also follows from compactness. □

Corollary 3.5. *Let $(X^{(p)}, X)$ and $(Y^{(p)}, Y)$ be point extension pairs and $h : X \to Y$ a surjective continuous map. There exists a continuous map $H : (X^{(p)}, X) \to (Y^{(p)}, Y)$ which restricts to h on X and to a homeomorphism of* $\mathrm{Iso}(X^{(p)}) = X^{(p)} \setminus X$ *onto* $\mathrm{Iso}(Y^{(p)}) = Y^{(p)} \setminus Y$. *In particular, if h is a homeomorphism, so is H.*

Proof. We attach Y to $X^{(p)}$ using h, letting $q_h : X^{(p)} \to X^{(p)}/h$ be the quotient map. We regard Y as a subset of $X^{(p)}/h$ so that $(X^{(p)}/h, Y)$ is a point extension pair and $q_h : X^{(p)} \to X^{(p)}/h$ is an extension of h which is a homeomorphism from $X^{(p)} \setminus X$ onto $X^{(p)}/h \setminus Y$. By Theorem 3.4 there is a homeomorphism $f : (X^{(p)}/h, Y) \to (Y^{(p)}, Y)$ rel Y. Let $H = f \circ q_h$. □

Lemma 3.6. *Suppose that $(X_1^{(p)}, X_1)$ and $(X_2^{(p)}, X_2)$ are point extension pairs and that $H : (X_1^{(p)}, X_1) \to (X_2^{(p)}, X_2)$ is a continuous map of pairs. Let (Y_1, X_1) and (Y_2, X_2) be extension pairs with $\pi_1 : (Y_1, X_1) \to (X_1^{(p)}, X_1)$ and $\pi_2 : (Y_2, X_2) \to (X_2^{(p)}, X_2)$ pair maps rel X_1 and rel X_2, respectively. Assume that $\widetilde{H} : Y_1 \to Y_2$ is a function such that $\pi_2 \circ \widetilde{H} = H \circ \pi_1$, i.e., the following diagram commutes:*

$$\begin{array}{ccc} Y_1 & \xrightarrow{\widetilde{H}} & Y_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ X_1^{(p)} & \xrightarrow[H]{} & X_2^{(p)} \end{array}$$

If for each $x \in \mathrm{Iso}(X_1^{(p)})$ the restriction $\widetilde{H} : \pi_1^{-1}(x) \to \pi_2^{-1}(H(x))$ is continuous, then $\widetilde{H}$ is continuous on Y_1.

Proof. If $x \in \mathrm{Iso}(X_1^{(p)})$, then $\pi_1^{-1}(x)$ is a clopen set, and thus $\widetilde{H}$ is continuous at the points of $\pi_1^{-1}(\mathrm{Iso}(X_1^{(p)}))$ by hypothesis.

Suppose $x \in X_1$, so $y = H(x) \in X_2$. Given $\epsilon > 0$, there exists by (3-1) $\delta > 0$ so that $\pi_2^{-1}(V_\delta^{d_2}(y)) \subset V_\epsilon^{d_2'}(y)$, where d_2 and d_2' are the metrics on $X_2^{(p)}$ and Y_2, respectively. By the continuity of H there exists $\gamma > 0$ so that $H(V_\gamma^{d_1}(x)) \subset V_\delta^{d_2}(y)$. If $x_1 \in \pi_1^{-1}(V_\gamma^{d_1}(x))$, then $\widetilde{H}(x_1) \in \pi_2^{-1}(V_\delta^{d_2}(y)) \subset V_\epsilon^{d_2'}(y)$. Hence, $\pi_1^{-1}(V_\gamma^{d_1}(x))$ is a neighborhood of x in Y_1 which is mapped by $\widetilde{H}$ into the neighborhood $V_\epsilon^{d_2'}(y)$ in Y_2, and continuity at x follows. □

Definition 3.7. For a space K and a pair of spaces (Y, X), we call Y a *K-extension* of X if there exist a point extension pair $(X^{(p)}, X)$ and a pair map $\pi : (Y, X) \to (X^{(p)}, X)$ rel X such that $\pi^{-1}(x)$ is homeomorphic to K for every $x \in \mathrm{Iso}(X^{(p)})$. We then call (Y, X) a *K-extension pair*, and the space Y is denoted by $X^{(K)}$.

We extend Lorch's theorem.

Theorem 3.8. *For any given space K, every space X has an essentially unique K-extension pair $(X^{(K)}, X)$. Furthermore, if $(X^{(K)}, X)$ and $(Y^{(K)}, Y)$ are K-extension pairs and $h : X \to Y$ is a surjective continuous map, then there exists a continuous map $H : (X^{(K)}, X) \to (Y^{(K)}, Y)$ which restricts to h on X and to a homeomorphism of $X^{(K)} \setminus X$ onto $Y^{(K)} \setminus Y$. In particular, if h is a homeomorphism, then so is H.*

Proof. Let $(X^{(p)}, X)$ be a point extension of X. Let $\pi : (X^{(p)} \times K, X \times K) \to (X^{(p)}, X)$ be the map of pairs given by the first coordinate projection. Attach $X^{(p)} \times K$ to X by using $h = \pi|_{X \times K}$. The map π factors through the quotient map q_h to define a map $((X^{(p)} \times K)/h, X) \to (X^{(p)}, X)$ rel X. Thus, $((X^{(p)} \times K)/h, X)$ is a K-extension pair.

Now let $\pi_X : (X^{(K)}, X) \to (X^{(p)}, X)$ and $\pi_Y : (Y^{(K)}, Y) \to (Y^{(p)}, Y)$ be maps rel X and Y, respectively, with each fiber over $X^{(p)} \setminus X$ and $Y^{(p)} \setminus Y$ homeomorphic to K. Use Corollary 3.5 to get an extension $H^{(p)} : (X^{(p)}, X) \to (Y^{(p)}, Y)$ which maps $X^{(p)} \setminus X$ to $Y^{(p)} \setminus Y$ homeomorphically. Define H by choosing for each $x \in X^{(p)} \setminus X$ an arbitrary homeomorphism from $\pi_X^{-1}(x)$ to $\pi_Y^{-1}(H^i(x))$. These exist because each fiber is homeomorphic to K. By Lemma 3.6, the resulting H is continuous.

In particular, if $Y = X$ and $h = 1_X$, then it follows that the K-extension is essentially unique. □

Two cases are of special interest to us. Recall that a component of a space X is an *isolated component* if it is a clopen subset of X.

Theorem 3.9. *Let (Y, X) be an extension pair.*

(a) *If K is a Cantor set, then (Y, X) is a K-extension pair if and only if Y is perfect and the dense, open set $Y \setminus X$ is zero-dimensional.*

(b) *If K is a connected space, then (Y, X) is a K-extension pair if and only if*

- *the union of the isolated components is dense in Y;*
- *each isolated component is homeomorphic to K;*
- (diameter condition) *for every $\epsilon > 0$ there are only finitely many isolated components with diameter greater than ϵ.*

Proof. Let $\{A_n : n \in \mathbb{N}\}$ be a pairwise disjoint sequence of nonempty clopen subsets of Y with union $Y \setminus X$. If for every $\epsilon > 0$ only finitely many of the sets A_n have diameter greater than ϵ, then

$$E = 1_X \cup \bigcup_n \{A_n \times A_n\} \tag{3-9}$$

is a closed equivalence relation. If $q : Y \to Y/E$ is the quotient space projection, then $(Y/E, X)$ is a point extension pair, q is a map of pairs rel X, and the fibers of q over the isolated points are the sets A_n. Conversely, if $\pi : (Y, X) \to (X^{(p)}, X)$ is a map of pairs rel X, then by (3-1) there are only finitely many fibers $\pi^{-1}(x)$ with diameter at least ϵ.

(a) If $\pi : (Y, X) \to (X^{(p)}, X)$ is a map rel X with each fiber over a point of $X^{(p)} \setminus X$ a Cantor set, then as a countable disjoint union of Cantor sets, $Y \setminus X$ is zero-dimensional and $\mathrm{Iso}(Y) = \varnothing$.

Conversely, if the locally compact space $Y \setminus X$ is zero-dimensional with no isolated points, then we can express it as the union of a pairwise disjoint sequence $\{C_n : n \in \mathbb{N}\}$ of nonempty, clopen subsets of Y each of which is thus a Cantor set. Let $\{C_{n,i} : i = 1, \ldots, N_n\}$ be a partition of C_n by nonempty clopen subsets of diameter less than n^{-1}. If $\{A_n\}$ is a counting of the collection $\{C_{n,i} : n \in \mathbb{N}, i = 1, \ldots, N_n\}$, then with E as in (3-9) the projection $q : (Y, X) \to (Y/E, X)$ is a map rel X with each fiber over a point of $(Y/E) \setminus X$ a Cantor set. Thus, (Y, X) is a Cantor set extension pair.

(b) If $\pi : (Y, X) \to (X^{(p)}, X)$ is a map rel X with each fiber connected, then $\{\pi^{-1}(x) : x \in X^{(p)} \setminus X\}$ is the set of isolated components, so the conditions of (b) are necessary; see Lemma 3.3.

Conversely, if they hold, then we let $\{A_n\}$ be the sequence of isolated components. There are infinitely many isolated components because X is nonempty and $Y \setminus X$ is dense. With E as in (3-9) again, $q : (Y, X) \to (Y/E, X)$ is the required map rel X. □

Corollary 3.10. *If Y and X are Cantor sets with closed, nowhere dense subsets $Y_1 \subset Y$ and $X_1 \subset X$, then for any surjective continuous map $h : Y_1 \to X_1$ there is a continuous map $H : Y \to X$ which extends h and which restricts to a homeomorphism of $Y \setminus Y_1$ to $X \setminus X_1$. In particular, if h is a homeomorphism, then so is H.*

Proof. By Theorem 3.9(a), (Y, Y_1) and (X, X_1) are Cantor set extension pairs. The existence of H then follows from Theorem 3.8. □

Remark. This is a classical theorem of Knaster and Reichbach [1953], extended by Gutek [1979].

Proposition 3.11. *Let $(X^{(p)}, X)$ be a point extension pair, $(X^{(C)}, X)$ a Cantor set extension pair and $(X^{(K)}, X)$ a K-extension pair for spaces X and K. If $\pi : (X^{(C)}, X) \to (X^{(p)}, X)$ is a surjective map of pairs rel X, then there exist $\pi_1 : (X^{(C)}, X) \to (X^{(K)}, X)$ and $\pi_2 : (X^{(K)}, X) \to (X^{(p)}, X)$ surjective maps of pairs rel X such that $\pi = \pi_2 \circ \pi_1$. That is, π factors through the pair $(X^{(K)}, X)$.*

Proof. For each $x \in \mathrm{Iso}(X^{(p)})$, $\pi^{-1}(x)$ is a clopen subset of $X^{(C)} \setminus C$ and so is compact, perfect and zero-dimensional, i.e., it is a Cantor set. By Proposition 2.9 there exists a continuous surjection h_x from $\pi^{-1}(x)$ onto K, so $E_x = (h_x \times h_x)^{-1}(1_K)$ is a closed equivalence relation on $\pi^{-1}(x)$ such that the surjection h_x factors to give a homeomorphism from the quotient space $\pi^{-1}(x)/E_x$ onto K. Let

$$E = 1_{X^{(C)}} \cup \bigcup_{x \in \mathrm{Iso}(X^{(p)})} E_x.$$

From Lemma 3.3 it follows that E is a closed equivalence relation on $X^{(C)}$. Let $q_E : X^{(C)} \to X^{(C)}/E$ be the quotient map. From the construction there exists $q : X^{(C)}/E \to X^{(p)}$ such that $\pi = q \circ q_E$. The E equivalence class of each $x \in X$ is a singleton and so, identifying X with $q_E(X)$, we can regard $q_E : (X^{(C)}, X) \to (X^{(C)}/E, X)$ and $q : (X^{(C)}/E, X) \to (X^{(p)}, X)$ as maps rel X. Since each $\pi^{-1}(x)/E_x$ is homeomorphic to K, $(X^{(C)}/E, X)$ is a K extension pair. By uniqueness, there exists $h : (X^{(C)}/E, X) \to (X^{(K)}, X)$ a homeomorphism rel X. Let $\pi_1 = h \circ q_E$ and $\pi_2 = q \circ h^{-1}$. □

Now we apply these results.

Lemma 3.12. *Let $(X^{(p)}, X)$ be a point extension pair, (Y, X) an extension pair and $\pi : (Y, X) \to (X^{(p)}, X)$ a surjective map of pairs rel X.*

(a) *The map $\pi : Y \to X^{(p)}$ is open.*

(b) *Assume that $H : (Y, X) \to (Y, X)$ and $H^{(p)} : (X^{(p)}, X) \to (X^{(p)}, X)$ are homeomorphisms with π mapping H to $H^{(p)}$, i.e., $\pi \circ H = H^i \circ \pi$. If $H^{(p)}$ is chain transitive, then H is chain transitive.*

Proof. (a) If $U \subset Y$ is open and $x \in \pi(U) \cap (X^{(p)} \setminus X)$, then $\pi(U)$ is a neighborhood of x because x is an isolated point.

If $x \in \pi(U) \cap X$, then there exist $\epsilon > 0$ and $\delta > 0$ so that $\overline{V}_{\epsilon}^{d_Y}(x) \subset U$ and $\pi^{-1}(\overline{V}_{\delta}^{d_{X^{(p)}}}(x)) \subset \overline{V}_{\epsilon}^{d_Y}(x)$. Because π is surjective, $\pi(\overline{V}_{\epsilon}^{d_Y}(x)) \supset \overline{V}_{\delta}^{d_{X^{(p)}}}(x)$, so $\pi(U)$ is a neighborhood of x.

(b) Since π is open, Proposition 2.4(e) implies that π is a semiconjugacy from $\mathcal{C}H$ to $\mathcal{C}H^{(p)}$ and from $\mathcal{C}H^{-1}$ to $\mathcal{C}(H^{(p)})^{-1}$. Fix $x \in X$. For any $y \in Y$ we have $x \in \mathcal{C}H^{(p)}(\pi(y))$ and $x \in \mathcal{C}(H^{(p)})^{-1}(\pi(y))$ because $H^{(p)}$ is chain transitive. Since $\{x\} = \pi^{-1}(x)$, it follows from the semiconjugacy that $x \in \mathcal{C}H(y)$ and $x \in \mathcal{C}H^{-1}(y)$. That is, every point of Y is chain equivalent to x, and so transitivity of $\mathcal{C}H$ implies that H is chain transitive. $\square$

Theorem 3.13. *Let $(X^{(p)}, X)$ be a point extension pair, $(X^{(C)}, X)$ a Cantor set extension pair and $(X^{(K)}, X)$ a K-extension pair for spaces X and K.*

(a) *If f is a chain transitive homeomorphism on X, then there exists a homeomorphism F on $X^{(p)}$ which extends f on X such that*
   - *F is chain transitive,*
   - *$F_{\pm}$ is topologically transitive, and*
   - *if $x \in X^{(p)} \setminus X$, then $\omega F(x) = X = \alpha F(x)$.*

(b) *There exists $G^{(C)}$ a topologically transitive homeomorphism on $X^{(C)}$ which extends 1_X.*

(c) *There exists $G^{(K)}$ a chain transitive homeomorphism on $X^{(K)}$ which extends 1_X.*

Proof. (a) By concatenating ϵ-chains which are ϵ-dense in X, we can obtain an infinite sequence $\{x_k : k \in \mathbb{Z}\}$ with $d(f(x_k), x_{k+1}) \to 0$ as $|k| \to \infty$ and so that for any $N \in \mathbb{N}$ the tails $\{x_k : k \geq N\}$ and $\{x_{-k} : k \geq N\}$ are dense in X. Define the sequence $\{y_k \in X \times [0, 1] : k \in \mathbb{Z}\}$ by

$$y_k = \begin{cases} (x_k, (2k+1)^{-1}) & \text{for } k \geq 0, \\ (x_k, (2|k|)^{-1}) & \text{for } k < 0. \end{cases} \tag{3-10}$$

Let $Y = X \times \{0\} \cup \{y_k : k \in \mathbb{Z}\}$, and define $\bar{F}(x, 0) = (f(x), 0)$ and $\bar{F}(y_k) = y_{k+1}$ for $k \in \mathbb{Z}$. It is easy to see that Y is an isolated point extension of $X = X \times \{0\}$, $\bar{F}$ is a homeomorphism on Y, and $X \times \{0\} = \omega\bar{F}(y_k) = \alpha\bar{F}(y_k)$ for any $k \in \mathbb{Z}$. Since the orbit $\mathcal{O}(\bar{F}_{\pm})(y_0) = \{y_k : k \in \mathbb{Z}\}$ is dense, it follows that $\bar{F}_{\pm}$ is topologically transitive. The homeomorphism $\bar{F}$ is chain transitive on Y because f is chain transitive on X and because $X = \omega\bar{F}(y_k) \subset \mathcal{C}\bar{F}(y_k)$ and also $X = \alpha\bar{F}(y_k) \subset \mathcal{C}\bar{F}^{-1}(y_k)$.

By Lorch's uniqueness theorem (Theorem 3.4) there exists a homeomorphism $H : (X^{(p)}, X) \to (Y, X)$ rel X. Let $F = H^{-1} \circ \bar{F} \circ H$.

(b) We begin with G a topologically transitive homeomorphism on a Cantor set with a Cantor set C of fixed points. By topological transitivity, C is necessarily nowhere dense. To be specific, let G be the shift homeomorphism on the product space $C^{\mathbb{Z}}$. Let $c : C \to C^{\mathbb{Z}}$ be the embedding with $c(x)_i = x$ for all $i \in \mathbb{Z}$. Thus, c is a homeomorphism onto the set of fixed points. Since C is nowhere dense, $(C^{\mathbb{Z}}, C)$ is a Cantor set extension pair by Theorem 3.9(a).

For an arbitrary space X, there exists by Proposition 2.9 a continuous surjection $h : C \to X$. We attach $C^{\mathbb{Z}}$ to X using h. Let $Y = (C^{\mathbb{Z}})/h$. Now, (Y, X) is a Cantor set extension pair, and G factors to define a topologically transitive homeomorphism $\overline{G}$ which restricts to the identity on X.

By Theorem 3.8 there exists a homeomorphism $H : (X^{(C)}, X) \to (Y, X)$ rel X. Let $G^{(C)} = H^{-1} \circ \overline{G} \circ H$.

(c) First we consider the case where X is the usual Cantor set C in $[0, 1]$ with $0, 1 \in C$. Then the complement consists of a pairwise disjoint, countable collection $\{(a_i, b_i) : i \in \mathbb{N}\}$ of open intervals in $(0, 1)$. Let $\ell_i = b_i - a_i$, so $\ell_i > 0$ for all $i \in \mathbb{N}$, but for any $\epsilon > 0$ there are only finitely many with $\ell_i \geq \epsilon$. Let

$$C^{(p)} = C \times \{0\} \cup \bigcup_{i,m \in \mathbb{N}} \{\langle a_i, \ell_i/m \rangle, \langle b_i, \ell_i/m \rangle\} \subset C \times [0, 1]. \tag{3-11}$$

Since the set of endpoints is dense in C, it follows that $C^{(p)}$ is a point extension of $C = C \times \{0\}$. Now we relabel the isolated points. For $i \in \mathbb{N}$, $k \in \mathbb{Z}$, define

$$\begin{aligned} u_{i,k} &= \begin{cases} \langle a_i, \ell_i \cdot (2k+1)^{-1} \rangle & \text{for } k \geq 0, \\ \langle b_i, \ell_i \cdot (2|k|)^{-1} \rangle & \text{for } k < 0, \end{cases} \\ v_{i,k} &= \begin{cases} \langle b_i, \ell_i \cdot (2k+1)^{-1} \rangle & \text{for } k \geq 0, \\ \langle a_i, \ell_i \cdot (2|k|)^{-1} \rangle & \text{for } k < 0. \end{cases} \end{aligned} \tag{3-12}$$

Define G as an extension of 1_C so that $u_{i,k} \xmapsto{G} u_{i,k+1}$ and $v_{i,k} \xmapsto{G} v_{i,k+1}$. Thus, above each endpoint a_i the $u_{i,k}$'s run up the $\langle a_i, \ell_i/m \rangle$'s with m even, jump from $\langle a_i, \ell_i/2 \rangle$ to $\langle b_i, \ell_i \rangle$, and then move down the $\langle b_i, \ell_i/m \rangle$'s with m odd. The $v_{i,k}$'s provide a similar path from b_i to a_i. Since for any $\epsilon > 0$ at most finitely many points move a distance more than ϵ, it follows that G and its inverse are continuous.

It is clear that for any i the points of $\{\langle a_i, \ell_i/m \rangle, \langle b_i, \ell_i/m \rangle : m \in \mathbb{N}\}$ all lie in a single chain component. Given $\epsilon > 0$, it is clear that we can get from a point $x \in C$ to a point $y \in C$ by an ϵ-chain jumping across the gaps of length less than ϵ which occur between x and y. For the finite number of remaining gaps we use the isolated point orbits to get across. Hence, G is chain transitive on $C^{(p)}$.

For an arbitrary space X, we again use a continuous surjection $h : C \to X$ from Proposition 2.9. Let $X^{(p)}$ be the quotient space $C^{(p)}/h$ obtained by attaching X via h. Then, $G \cup 1_X$ on $C^{(p)} \cup X$ factors through the quotient map q_h to define a homeomorphism $G^{(p)}$. Because q_h maps G on $C^{(p)}$ onto $G^{(p)}$ on $X^{(p)}$ it follows that $G^{(p)}$ is chain transitive by Proposition 2.4(d). Because $q_h : C^{(p)} \setminus C \to X^{(p)} \setminus X$ is a homeomorphism, it follows that $X^{(p)}$ is an isolated point extension of X.

Because $(X^{(K)}, X)$ is a K-extension pair and the point extension is essentially unique, there exists $\pi : (X^{(K)}, X) \to (X^{(p)}, X)$, a map of pairs rel X, such that the fiber $\pi^{-1}(x)$ is homeomorphic to K for every $x \in X^{(p)} \setminus X$. For each such x let $G^{(K)}$ restrict to a homeomorphism from $\pi^{-1}(x)$ to $\pi^{-1}(G^{(p)}(x))$. By Lemma 3.6 $G^{(K)}$

and its inverse are continuous. By Lemma 3.12 $G^{(K)}$ is chain transitive because $G^{(p)}$ is. □

Remark. By Proposition 2.5(b) there is no topologically transitive homeomorphism on a space with infinitely many isolated points. Hence, the result in (a) above is the best we can hope for. In particular, we see that any chain transitive homeomorphism on X can be extended to a system in which X is an omega limit set.

Recall from Proposition 2.7(e) that if X is connected, then 1_X is chain transitive. Now we can prove a slight extension of Theorem 1.1.

Corollary 3.14. *For a space X let X_1 be the closure of the union of all components which meet $\overline{\mathrm{Iso}(X)}$. If X_1 is a proper, clopen, nonempty subset of X, then X is $H(X)$-decomposable and so X admits no chain transitive homeomorphism. If X_1 is not a proper, clopen subset of X and the open set $X \setminus X_1$ is empty or zero-dimensional, then X admits a chain transitive homeomorphism.*

Proof. The sets $\overline{\mathrm{Iso}(X)}$ and X_1 are h_X-invariant, so if X_1 is proper, clopen and nonempty, then X is h_X-decomposable.

Assume that X_1 is not a proper, clopen subset of X. If $\mathrm{Iso}(X)$ is finite and nonempty, then $X = X_1 = \mathrm{Iso}(X)$ and we can define f so that X consists of a single periodic orbit. If $\mathrm{Iso}(X) = \varnothing$, then $X_1 = \varnothing$ and $X = X \setminus X_1$ is zero-dimensional and perfect, and so X is a Cantor set. Hence, X admits a topologically transitive homeomorphism.

Now assume that $\mathrm{Iso}(X)$ is infinite, so $A = \overline{\mathrm{Iso}(X)} \setminus \mathrm{Iso}(X)$ is nonempty. Clearly, $(\overline{\mathrm{Iso}(X)}, A)$ is a point extension pair, and by Theorem 3.13(c) there exists a chain transitive homeomorphism f_1 on $\overline{\mathrm{Iso}(X)}$ which is the identity on A. Extend f_1 to be the identity on $X_1 \setminus \mathrm{Iso}(X)$. Thus, f_1 is the identity on every nontrivial component of X which meets $\overline{\mathrm{Iso}(X)}$. It follows from Proposition 2.7(e) that all of these are contained in the chain component of f_1 which contains all of $\overline{\mathrm{Iso}(X)}$. As this chain component is closed, it must contain all of X_1. That is, f_1 on X_1 is chain transitive. If $X = X_1$, then we are done.

Otherwise, the nonempty, open, zero-dimensional set $X \setminus X_1$ is not closed and contains no isolated points. So $X_2 = \overline{X \setminus X_1}$ is perfect and $B = X_2 \cap X_1$ is nonempty subset of X_1 disjoint from $\mathrm{Iso}(X)$. We see that (X_2, B) is a Cantor set extension pair, so by Theorem 3.13(b) there exists a topologically transitive homeomorphism f_2 on X_2 which restricts to the identity on B.

The concatenation $f = f_1 \cup f_2$ is a homeomorphism on X. Since f_1 and f_2 are each chain transitive and $X_1 \cap X_2 \neq \varnothing$ it follows that all of X is contained in a single chain component, i.e., f is chain transitive. □

To extend these results we need some simple lifting facts.

Lemma 3.15. *Let $f_i \in H(X_i)$ for $i = 1, 2$ and let $\pi : X_1 \to X_2$ be a continuous surjection mapping f_1 to f_2. Assume that $(\pi \times \pi)^{-1}(1_{X_2}) \subset \mathcal{C}f_1$. That is, each fiber of π is entirely contained in a single chain component of f_1.*

(a) *Both f_1 and f_2 are chain recurrent, i.e.,*

$$1_{X_1} \subset \mathcal{C}f_1 \quad \text{and} \quad 1_{X_2} \subset \mathcal{C}f_2.$$

(b) *The space X_1 is f_1-decomposable if and only if X_2 is f_2-decomposable.*

(c) *The chain relations satisfy*

$$\mathcal{C}f_2 = (\pi \times \pi)(\mathcal{C}f_1) \quad \text{and} \quad \mathcal{C}f_1 = (\pi \times \pi)^{-1}(\mathcal{C}f_2).$$

(d) *The homeomorphism f_1 is chain transitive if and only if f_2 is chain transitive.*

Proof. (a) Let $E_\pi = (\pi \times \pi)^{-1}(1_{X_2})$. It is a closed equivalence relation and so contains 1_{X_1}. Hence, $1_{X_1} \subset E_\pi \subset \mathcal{C}f_1$.

Because π is surjective, $1_{X_2} = (\pi \times \pi)(1_{X_1})$. Since π maps f_1 to f_2, it maps $\mathcal{C}f_1$ to $\mathcal{C}f_2$ by Proposition 2.4(d). Hence,

$$1_{X_2} = (\pi \times \pi)(1_{X_1}) \subset (\pi \times \pi)(\mathcal{C}f_1) \subset \mathcal{C}f_2.$$

(b) If B and its complement are proper, clopen, forward f_2-invariant subsets of X_2 then because π is surjective, $\pi^{-1}(B)$ and its complement are proper, clopen, forward f_1-invariant subsets of X_1.

Now assume that A and its complement are proper, clopen, forward f_1-invariant subsets of X_1. Since each is forward $\mathcal{C}f_1$-invariant, it follows that each is saturated by the equivalence relation $E_\pi \subset \mathcal{C}f_1$. Hence, $\pi(A)$ and $\pi(X_1 \setminus A)$ are disjoint closed sets with union $\pi(X_1) = X_2$. That is, they are complementary clopen sets. Furthermore, they are forward f_2-invariant.

(c) As mentioned above, $(\pi \times \pi)(\mathcal{C}f_1) \subset \mathcal{C}f_2$. Now assume $(x_1, x_2) \notin \mathcal{C}f_1$. Since $X_1 = |\mathcal{C}f_1|$ by (a), Proposition 2.1(c) implies there is an attractor A for f_1 which contains x_1 but not x_2, and by Proposition 2.1(d) A is a clopen $\mathcal{C}f_1$-invariant set. Hence, it is saturated by E_π. So $\pi(A)$ is clopen and f_2-invariant, and therefore also $\mathcal{C}f_2$-invariant. Now, $\pi(x_1) \in \pi(A)$, and $x_2 \notin A = \pi^{-1}(\pi(A))$ implies $\pi(x_2) \notin \pi(A)$. It follows that $(\pi(x_1), \pi(x_2)) \notin \mathcal{C}f_2$. Thus, the complement of $\mathcal{C}f_1$ in $X_1 \times X_1$ maps into the complement of $\mathcal{C}f_2$. Since $\pi \times \pi$ is surjective, the equations of (c) follow.

(d) Immediate from (c) and the surjectivity of π. □

For any space X, the chain relation $\mathcal{C}1_X$ is a closed equivalence relation and by Lemma 2.6(b) the equivalence classes are the components of X. Let $[X]$ be the zero-dimensional space of components, the quotient space for this equivalence relation with quotient map $\pi_X : X \to [X]$; see Lemma 2.6(b) and Proposition 2.1(d). There is a natural homomorphism $[\cdot] : H(X) \to H([X])$ with π_X mapping f to

$[f]$ for $f \in H(X)$. Thus, $H(X)$ acts on $[X]$ and we let $[h_X] = \bigcup\{[h] : h \in H(X)\}$ be the associated relation on $[X]$.

Proposition 3.16.

(a) *A space X is h_X-decomposable if and only if $[X]$ is $[h_X]$-decomposable.*

(b) *If f is a chain recurrent homeomorphism on a space X, then $(\pi_X \times \pi_X)^{-1}(1_{[X]}) \subset \mathcal{C}f$. In that case, the following are equivalent:*

(i) *The map f is chain transitive.*

(ii) *The space X is f-indecomposable.*

(iii) *The map $[f]$ is chain transitive.*

(iv) *The space $[X]$ is $[f]$-indecomposable.*

Proof. (a) This is clear because any clopen set is saturated by the equivalence relation $\mathcal{C}1_X$.

(b) Since the space of chain components is zero-dimensional, every connected set of chain recurrent points is contained in a single chain component. If f is chain recurrent, then every point is chain recurrent, so each component is contained in a single chain component. It follows that $(\pi_X \times \pi_X)^{-1}(1_{[X]}) \subset \mathcal{C}f$. Since f and $[f]$ are chain recurrent, the equivalences (i) $\Leftrightarrow$ (ii) and (iii) $\Leftrightarrow$ (iv) follow from Proposition 2.7(d). The implication (i) $\Rightarrow$ (iii) holds because π maps f to $[f]$. The converse, (iii) $\Rightarrow$ (i), follows from Lemma 3.15. $\square$

Thus, a chain transitive homeomorphism on $[X]$ lifts to a chain transitive homeomorphism if and only if it lifts to a chain recurrent homeomorphism.

Recall that a component K of X is an *isolated component* if it is a clopen subset of X. For a space X let $\mathcal{I}_X$ denote the set of isolated components. Two isolated components K_1 and K_2 are $H(X)$-*equivalent* if they are homeomorphic or, equivalently, if there exists $g \in H(X)$ such that $g(K_1) = K_2$. Let I_X be the set of $H(X)$-equivalence classes in $\mathcal{I}_X$. For $i \in I_X$ let Q_i be the union of the isolated components in i, and let Q be the union of all of the isolated components, so that Q is the disjoint union of the Q_i. Thus, Q and all of the Q_i are open subsets of X.

Lemma 3.17. *If A is a clopen $H(X)$-invariant set which meets some $\overline{Q_i}$, then it contains $\overline{Q_i}$. In particular, if all the isolated components are homeomorphic to one another and the union of the isolated components is dense, then X is $H(X)$-indecomposable.*

Proof. Since A is open, it meets some isolated component $K \in i$. Since it is clopen, it contains K. If $K_1 \in Q_i$, then there exists $g \in H(X)$ such that $g(K) = K_1$, and so $H(X)$-invariance implies $K_1 \subset A$. Since $Q_i \subset A$ and A is closed, we get $\overline{Q_i} \subset A$. $\square$

We say that X satisfies the *diameter condition on isolated components* if for every $\epsilon > 0$ there are only finitely many isolated components with diameter greater than ϵ.

The following is the furthest we can extend Corollary 3.14.

Theorem 3.18. *For a space X, let X_1 be the closure of the union of all components which meet the closure of the union of all isolated components. If X satisfies the diameter condition on isolated components and the open set $X \setminus X_1$ is empty or zero-dimensional, then*

(a) *either X is $H(X)$-decomposable or X admits a chain transitive homeomorphism;*

(b) *if X_1 is a proper, clopen subset of X, then X is $H(X)$-decomposable, and so X admits no chain transitive homeomorphism;*

(c) *if X_1 is not a proper, clopen subset of X and all the isolated components are homeomorphic to one another, then X admits a chain transitive homeomorphism.*

Proof. (b) Obvious since X_1 is $H(X)$-invariant.

In (a) and (c) the extension from the closure of the isolated components to the rest of X proceeds just as in Corollary 3.14. So from now on we will assume that the union Q of the isolated components is dense.

(c) If there are only n isolated components, then their union Q is clopen and so is all of X. By assumption they are all homeomorphic to some common space K. We can choose a homeomorphism with $f^n = 1_X$, and so that each periodic orbit meets each component. Clearly, f is chain transitive.

Now we may assume that $\mathcal{I}_X$ is infinite, and let $A = X \setminus Q$. Thus, A is a nonempty, closed, nowhere dense set. By Theorem 3.9(b) (X, A) is a K-extension pair. By Theorem 3.13(c) X admits a chain transitive homeomorphism rel A.

(a) We may assume that there is more than one equivalence class in I_X, for otherwise we are in case (c). If any $i \in I_X$ is finite, then Q_i is a clopen $H(X)$-invariant set, so X is $H(X)$-decomposable since $X \neq Q_i$ by the assumption that I_X contains more than one class.

Now assume that every equivalence class in I_X is infinite. Then the closure of each open $H(X)$-invariant set Q_i meets A, and we let $A_i = \overline{Q_i} \cap A$. If $i \neq j$, then $\overline{Q_i} \cap \overline{Q_j} \subset A$, and so this intersection equals $A_i \cap A_j$.

Applying the argument for (c) to $\overline{Q_i}$, there exists a homeomorphism f_i on $\overline{Q_i}$ which is chain transitive and which restricts to the identity on A_i.

Let f on X equal f_i on $\overline{Q_i}$ and the identity on A. Because the diameter condition holds, we can apply Lemma 3.12 to see that f is a homeomorphism on X. Since each f_i is chain transitive, $\overline{\bigcup_i \{Q_i \times Q_i\}} \subset \mathcal{C}f$.

Let $x \in X$ and $g \in H(X)$. Since $Q = \bigcup_i Q_i$ is dense in X, there is a sequence $\{x_k \in Q_{i_k}\}$ which converges to x. Then $g(x_k) \in Q_{i_k}$, and so $(x, g(x)) \in \overline{\bigcup_i Q_i \times Q_i}$. Hence, $h_X \subset \mathcal{C}f$. In particular, $1_X \subset h_X$ implies that f is chain recurrent. So by Proposition 2.7(d) f is chain transitive if and only if X is f-indecomposable. Since $h_X \subset \mathcal{C}f$, an f-decomposition, which is a $\mathcal{C}f$-decomposition, is also an h_X-decomposition. Hence, if X is f-indecomposable, then it is h_X-indecomposable, i.e., X is $H(X)$-indecomposable. On the other hand, if X is $H(X)$-decomposable, then there does not exist any chain transitive homeomorphism on X. $\square$

As we will see below, the diameter condition on isolated components is essential for this result.

Example 3.19. We construct X so that

- the connected components of X are all homeomorphic to $[0, 1]$;
- the space X is $H(X)$-decomposable;
- there are no isolated components.

Let $C \subset [0, 1]$ be a Cantor set and $S = \{a_n : n \in \mathbb{N}\}$ a sequence of distinct points in C with closure A in C. Define

$$I_n = \{\langle a_n, t\rangle : 0 \le t \le n^{-1}\} \subset C \times [0, 1] \quad \text{for } n \in \mathbb{N},$$

$$C_0 = C \times \{0\}, \quad C_+ = C_0 \cup \bigcup_n I_n, \quad C_\pm = C_+ \cup C \times [-1, 0], \tag{3-13}$$

$$A_0 = A \times \{0\}, \quad A_+ = \overline{\bigcup_n I_n}, \quad A_\pm = A_+ \cup A \times [-1, 0].$$

Each $I_n^\circ = I_n \setminus C \times \{0\}$ is open in $C_\pm$, and so the points of each I_n° have connected neighborhoods. Hence, A, A_+ and $A_\pm$ are $H(C_\pm)$-invariant. Thus, if A is a proper, clopen subset of C, then $C_\pm$ is $H(C_\pm)$-decomposable. Observe that every component is homeomorphic to the unit interval and the first coordinate projection maps $C_\pm$ onto the Cantor set C.

Also, if A is a proper, clopen set, then $C \times [-1, 0]$ admits chain transitive homeomorphisms, but the factor $X = C \cup A \times [-1, 0]$ is $H(X)$-decomposable.

A homeomorphism f on C can be extended to a homeomorphism F_+ of C_+ if and only if A is f-invariant. In that case, we can then define F_+ by using any orientation-preserving homeomorphism from I_x to $I_{f(x)}$, i.e., one which maps $\langle x, 0\rangle$ to $\langle f(x), 0\rangle$. Here $I_x = I_n$ for $x = a_n$ and $= \{(x, 0)\}$ if $x \notin S$. Continuity at points of A_+° is clear, and if $x \in C$, then for every ϵ there is a neighborhood U of x in C so that $y \in U \setminus \{x\}$ implies that the length of the interval $I_{f(y)}$ is less than ϵ. This implies continuity at $(x, 0)$. Notice that if $\mathcal{O}f(x)$ is infinite, then

$$\lim_{|n|\to\infty} |I_{f^n(x)}| = 0, \tag{3-14}$$

where $|J|$ denotes the length of an interval J. This says that any pair $(x, t_1), (x, t_2) \in I_x$ is *asymptotic* for F_+ and $(F_+)^{-1}$ with

$$\begin{aligned} \omega F_+(x, t_1) = \omega F_+(x, t_2) = \omega f(x) \times \{0\} \subset C_0, \\ \alpha F_+(x, t_1) = \alpha F_+(x, t_2) = \alpha f(x) \times \{0\} \subset C_0. \end{aligned} \tag{3-15}$$

Now assume that A is not clopen. If f is chain transitive and every point of A has an infinite orbit, then any extension F_+ to C_+ is chain transitive. Observe that F_+ is chain transitive if, whenever a_n is a periodic point for f, we define F_+ by using the unique linear, orientation-preserving homeomorphism from I_{a_n} to $I_{f(a_n)}$. On the other hand, if $f(a_1) = a_1$ and on I_{a_1} we define F_+ by $(a_1, t) \mapsto (a_1, t^2)$ then $(a_1, 1)$ is a repelling fixed point for F_+ on I_{a_1} and hence for F_+ on C_+. If F_+ is chain transitive, then we can obtain a chain transitive extension $F_\pm$ on $C_\pm$ by using $f \times 1_{[-1,0]}$ on $C \times [-1, 0]$. □

Example 3.20. We construct a space X so that

- the connected components of X are all homeomorphic to $[0, 1]$;
- the space of connected components, $[X]$, consists of a convergent sequence and its limit, and the union of isolated components in X is dense;
- it is f-decomposable for any $f \in H(X)$ but not $H(X)$-decomposable.

The space X we construct is $H(X)$-indecomposable by the first two properties and by Lemma 3.17.

For every $n \in \mathbb{N}$ we define $I_n = [0, n^{-1}]$ and a continuous function $t_n : I_n \to I = [0, 1]$ so that, for integers $i = 0, \ldots, (2n)!$,

$$t_n\left(\frac{i}{n(2n)!}\right) = \begin{cases} 0 & \text{when } i \text{ is even,} \\ 1 & \text{when } i \text{ is odd,} \end{cases}$$

and the rest of the values are defined by linear interpolations.

Each interval I_n contains $(2n)!$ intervals $\{I_n^i : i = 1, \ldots, (2n)!\}$ of equal length, each of which is mapped by t_n onto I. We can further subdivide each I_n^i into intervals $\{I_n^{i,j} : j = 1, \ldots, n\}$ of equal length so that each is mapped to a subinterval of I of length n^{-1} by t_n. The corresponding restrictions of t_n are denoted by $t_{n,i} = t_n|_{I_n^i}$ and $t_{n,i,j} = t_n|_{I_n^{i,j}}$.

Recall that we identify a function with its graph, so the functions t_n, $t_{n,i}$ and $t_{n,i,j}$ are all closed subsets of $I_n \times I$. We define

$$\begin{aligned} X_n &= \{n^{-1}\} \times t_n, \quad n \in \mathbb{N}, \\ X_\infty &= \{\langle 0, 0\rangle\} \times [0, 1], \\ X &= \bigcup_n X_n \cup X_\infty. \end{aligned} \tag{3-16}$$

The space X is clearly a closed, bounded subset of $\mathbb{R}^3$, and the space $[X]$ can be identified with $\pi(X)$, where $\pi : X \to \{n^{-1} : n \in \mathbb{N}\} \cup \{0\}$ is the projection to the first coordinate. The union of the isolated components is clearly dense in X. In addition, for all appropriate n, i and j, we define

$$X_n^i = \{n^{-1}\} \times t_{n,i} \quad \text{and} \quad X_n^{i,j} = \{n^{-1}\} \times t_{n,i,j}. \tag{3-17}$$

Note that each X_n is a union of the line segments X_n^i. Similarly, each X_n^i is a union of the line segments $X_n^{i,j}$. The diameter of each X_n^i is greater than 1, and the diameter of each $X_n^{i,j}$ is less than $2n^{-1}$. The arc length of X_n is greater than $(2n)!$.

Suppose that $h : X_n \to X_m$ is a homeomorphism for some $n < m < \infty$. Then for some $i = 1, \ldots, (2n)!$, $j = 1, \ldots, n$ and $k = 1, \ldots, (2m)!$, it must happen that $h(X_n^{i,j}) \supset X_m^k$. If not, then each $h(X_n^{i,j})$ meets at most two of the segments X_m^i, so the arc length of each mapped segment $h(X_n^{i,j})$ is less than 4. The arc length of $h(X_n)$, which is the sum of the arc lengths of the $(2n)!n$ mapped segments $h(X_n^{i,j})$, is less than $4(2n)!n$. Since $4(2n)!n < (2m)!$, the map h cannot be surjective. It follows that there exists a pair of points $u, v \in X_n$ with $d(u, v) \le 2n^{-1}$ but with $d(h(u), h(v)) \ge 1$.

Now if $f \in H(X)$, then it follows that the induced homeomorphism $[f]$ on the space $[X]$ is the identity on all but finitely many points. For if not, then by replacing f by f^{-1} if necessary, we can assume that there are sequences (m_i) and (n_i) in $\mathbb{N}$ tending to infinity such that $f(X_{n_i}) = X_{m_i}$ and $m_i > n_i$ for all i. Hence, there exist $u_i, v_i \in X_{n_i}$ with $d(u_i, v_i) \le 2n_i^{-1}$ but with $d(f(u_i), f(v_i)) \ge 1$. Thus, there are convergent subsequences of $\{u_i\}$ and $\{v_i\}$ with a common limit in X_∞. Hence, f cannot extend to a continuous function on all of X.

Thus, there exists $N \in \mathbb{N}$ such that $f(X_n) = X_n$ for all $N \le n \le \infty$. Since the isolated components X_n are invariant for n large enough, X is f-decomposable. □

Example 3.21. We construct spaces X and X^+ so that

- the isolated components of X and X^+ are all homeomorphic to one another and their union is dense;
- X^+ is $H(X^+)$-indecomposable but it is f-decomposable for all $f \in H(X^+)$;
- there exists $f \in H(X)$ such that X is f-indecomposable, but no $f \in H(X)$ is chain transitive.

Let $Z = \{x_n : n \in \mathbb{Z}\} \subset I = [0, 1]$ with

$$x_n = \begin{cases} 1 - (n+2)^{-1} & \text{for } n = 0, 1, \ldots, \\ (|n|+2)^{-1} & \text{for } n = -1, -2, \ldots. \end{cases}$$

Define for $n \in \mathbb{Z}$

$$\begin{gathered} a_n = \langle x_n, 0\rangle, \quad b_n = \langle x_n, (|n|+1)^{-1}\rangle, \\ I_n = \{\langle x_n, t\rangle : 0 \le t \le (|n|+1)^{-1}\}. \end{gathered} \tag{3-18}$$

Define

$$\begin{gathered} J = I \times \{0\}, \quad H = [-1, 0] \times \{0\}, \\ C = H \cup J \cup \bigcup \{I_n : n \in \mathbb{Z}\}, \\ Z_0 = \{a_n : n \in \mathbb{Z}\}, \quad Z_1 = \{b_n : n \in \mathbb{Z}\}, \\ e_0 = \langle 0, 0\rangle, \quad e_1 = \langle 1, 0\rangle, \quad e_{-1} = \langle -1, 0\rangle. \end{gathered} \tag{3-19}$$

We will call C a *comb* with handle H.

The group $H(C)$ fixes e_{-1}, e_0 and e_1. Each of the sets Z_0 and Z_1 is a single $H(C)$-orbit. The closed sets J and H are $H(C)$-invariant.

If $h \in H(C)$, then for some k

$$h(b_n) = b_{n+k} \quad \Leftrightarrow \quad h(a_n) = a_{n+k} \quad \Leftrightarrow \quad h(I_n) = I_{n+k}.$$

In that case $h(a_{n\pm 1}) = a_{n\pm 1+k}$, and so h induces a translation by k on the sequences Z_0 and Z_1. If $k > 0$, then e_1 is an attractor with complementary repeller H, and the reverse is true if $k < 0$. If $k = 0$, then h fixes each point of Z_0 and of Z_1.

On C define T by

$$\begin{aligned} T(e_0) &= e_0, & T(e_{\pm 1}) &= e_{\pm 1}, \\ T(a_n) &= a_{n+1}, & T(b_n) &= b_{n+1}, \end{aligned} \tag{3-20}$$

and with $T : [a_{n-1}, a_n] \to [a_n, a_{n+1}]$, $T : I_n \to I_{n+1}$ and $T : H \to H$ linear for all $n \in \mathbb{Z}$. Thus, $T \in H(C)$ induces a translation by 1 on Z_0 and Z_1 and is the identity on H.

For $n \in \mathbb{Z}$ let

$$\begin{aligned} I_n^* &= I_n \cup \big\{(x, (|n|+1)^{-1}) : \tfrac{1}{2}(x_{n-1} + x_n) \le x \le \tfrac{1}{2}(x_{n+1} + x_n)\big\}, \\ C_n &= C \cup I_n^*, \end{aligned} \tag{3-21}$$

a comb with a queer tooth I_n^* at n replacing I_n. Observe that C and all the C_n are connected.

If $f : C_n \to C_m$ is a homeomorphism, then $f(I_n^*) = I_m^*$. If $h \in H(C)$, then h extends to a homeomorphism from C_n to C_m if and only if $h(I_n) = I_m$. Thus, C_n and C_m are homeomorphic for all $n, m \in \mathbb{Z}$.

In $\mathbb{R}^2 \times I$ we define

$$\begin{aligned} X &= C \times \{0, 1\} \cup \bigcup \{C_n \times \{x_n\} : n \in \mathbb{Z}\}, \\ X^+ &= C \times \{1\} \cup \bigcup \{C_n \times \{x_n\} : n \in \mathbb{N}\}, \end{aligned} \tag{3-22}$$

Thus, X and X^+ both have a dense union of isolated components, and each isolated component is homeomorphic to C_0. It follows from Lemma 3.17 that X is $H(X)$-indecomposable and X^+ is $H(X^+)$-indecomposable.

Any homeomorphism in $H(X^+)$ restricts to a homeomorphism of $C \times \{1\}$ and so restricts to translation by k on $Z_0 \times \{1\}$. This means that h must map $C_n \times \{x_n\}$ to $C_{n+k} \times \{x_{n+k}\}$ when $n \ge N$ for N sufficiently large, and we may suppose $N + k > 0$. This implies that h maps the set of N complementary components $\{C_n \times \{x_n\} : 0 \le n < N\}$ to the set of $N+k$ components $\{C_n \times \{x_n\} : 0 \le n < N+k\}$. This requires that $k = 0$. Hence, each isolated component $C_n \times \{x_n\}$ with $n \ge N$ is invariant, and so X^+ is h-decomposable.

Thus, X^+ is an example of a space which is $H(X^+)$-indecomposable but such that X^+ is h-decomposable for every $h \in H(X^+)$.

In the case of X, we start by assuming that h fixes each of the two nonisolated components, i.e., h maps $C \times \{\epsilon\}$ to itself for $\epsilon = 0, 1$. As before, since h translates $Z_0 \times \{\epsilon\}$ by some k_ϵ for $\epsilon = 0, 1$, there exists $N \in \mathbb{N}$ large enough that h maps $C_n \times \{x_n\}$ to $C_{n+k_1} \times \{x_{n+k_1}\}$ for $n \geq N$ and $C_n \times \{x_n\}$ to $C_{n+k_0} \times \{x_{n+k_0}\}$ for $-n \geq N$. Again we can choose N large enough that $N + k_\epsilon > 0$ for $\epsilon = 0, 1$. This implies that h maps the set of complementary components $C_n \times \{x_n\}$ for $-N < n < N$ to the set of components $C_n \times \{x_n\}$ for $-N + k_0 < n < N + k_1$. This requires that $k_0 = k_1$, and we let $k = k_0 = k_1$.

If $k > 0$, then $C \times \{1\}$ is an attractor and $C \times \{0\}$ is a repeller with the reverse if $k < 0$. Hence, h is not chain transitive. If $k = 0$, then each isolated component $C_n \times \{x_n\}$ is invariant for $|n| \geq N$, and so X is h-decomposable.

The remaining possibility is that h interchanges the two limit components $C \times \{\epsilon\}$ for $\epsilon = 0, 1$ and then each is invariant for h^2. Applying the previous argument to h^2, we see that for a large N the components are translated by h^2 with a common k. If $k > 0$, then $C \times \{1\}$ is an attractor for h^2 while $C \times \{0\}$ is a repeller. But since h commutes with h^2 this would imply that $C \times \{0\} = h(C \times \{1\})$ would be an attractor for h^2 as well, which it is not. Similarly, $k < 0$ leads to a contradiction. It follows that in this interchange case $k = 0$, and so $C_n \times \{x_n\}$ is h^2-invariant for $|n|$ sufficiently large. Hence, for each such n the set $C_n \times \{x_n\} \cup h(C_n \times \{x_n\})$ is clopen and h-invariant, so X is h-decomposable when h interchanges the ends.

Finally, extend $T : C \to C$ by

$$T_n : C_n \times \{x_n\} \to C_{n+1} \times \{x_{n+1}\}$$

for all $n \in \mathbb{Z}$ and by $T \times 1_{\{0,1\}}$ on $C \times \{0, 1\}$ to obtain a homeomorphism (with $k = 1$) with respect to which X is not decomposable.

Thus, X is an example of a space with an element $h \in H(X)$ such that X is not h-decomposable, but nonetheless there is no chain transitive homeomorphism in $H(X)$. □

4. Spaces with all homeomorphisms chain transitive

Having considered spaces which admit no chain transitive homeomorphisms, we turn to the opposite extreme to consider spaces such that every homeomorphism is chain transitive. By Proposition 2.7(e) the identity 1_X is chain transitive if and only if X is connected, and if $f \in H(X)$ with X connected, then f is chain transitive if and only if it is chain recurrent.

A space X is called *rigid* if 1_X is the only homeomorphism on X, i.e., the group $H(X)$ is trivial. Such spaces were introduced and constructed by de Groot and Wille

[1958]. For a connected rigid space the only element of $H(X)$ is chain transitive. We construct some more interesting examples by using rigid spaces as tools. We need a pairwise disjoint sequence $\{Z_n\}$ of connected, locally connected spaces such that

(RIG) for any nonempty open subset U of Z_n and any disk I^k, there does not exist a homeomorphism of $U \times I^k$ onto any subset of $(Z_n \setminus U) \times I^k$ or onto any subset of $Z_m \times I^k$ with $m \neq n$;

(CON) for every $n \in \mathbb{N}$ and for any finite $F \subset Z_n$, the set $Z_n \setminus F$ has only finitely many components, and for any positive integer N there is a subset $F \subset Z_n$ of cardinality N such that $Z_n \setminus F$ is connected.

Condition RIG is a slight strengthening of the condition on a space called *strongly chaotic* in [Charatonik and Charatonik 1996]. We construct such a sequence in the Appendix. For each Z_n we choose a pair of distinct points $e_n^-, e_n^+ \in Z_n$. Our examples are obtained by using such rigid spaces instead of the unit interval in some common constructions.

Example 4.1. Suppose that G is a finitely generated (and hence countable) group. Following de Groot [1959], we construct a space X so that

- the homeomorphism group $H(X)$ is isomorphic to G;
- every $f \in H(X)$ is chain transitive.

First, consider the case $G = \mathbb{Z}$. We can think of the real line as a graph with $\mathbb{Z}$ as the set of vertices and intervals $[n, n+1]$ as edges. Now let $Z = Z_1$ be one of the chaotic spaces described above with points $e^-, e^+ \in Z$. We replace each edge by a copy of Z. That is, let X_0 be the quotient space of $\mathbb{Z} \times Z$ with (n, e^+) identified with $(n+1, e^-)$. If t is the translation homeomorphism on $\mathbb{Z}$ with $t(n) = n+1$, then t has a unique extension t to X_0 which is the quotient of $t \times 1_Z$. The only homeomorphisms on X_0 are the iterates t^n. Let X be the one-point compactification of X_0 with the additional point ∞. Let $t \in H(X)$ be the unique homeomorphism extension of t on X_0 and so of t on $\mathbb{Z}$. Since X is connected, 1_X is chain transitive. For any $n \neq 0$, we have $\{\infty\} = \omega(t^n)(x) = \alpha(t^n)(x)$ for all $x \in X$, and so t^n is chain transitive. In this case, $H(X)$ is isomorphic to $\mathbb{Z}$.

In general, suppose that G is generated by $\{g_1, \dots, g_n\}$. Let X_0 be the Cayley graph with rigid spaces as linking edges. That is, let $\{g_1, \dots, g_n\}$ be a list of generators for G, and let $\{Z_1, \dots, Z_n\}$ be distinct strongly chaotic spaces as above, each with a chosen pair of points. Let X_0 be the quotient space of $G \cup \big[G \times \big(\bigcup_{i=1}^n Z_i\big)\big]$ with (g, e_i^+) identified with $(g_i g, e_i^-)$ for $g \in G$, $i = 1, \dots, n$ and (g, e_i^-) identified with $g \in G$ for $i = 1, \dots, n$. Because there are only finitely many generators, the space X_0 is locally compact and the set of vertices $\{g \in G\}$ is invariant with respect to any homeomorphism h. Furthermore, if $g \in G$, then $h(g_i g) = g_i h(g)$, and so h commutes with all left translations. It follows that, on G, the mapping h is the right translation r_v, where $v = h(u)$ and u is the identity element of G. Thus, on X_0, the mapping h is the quotient of the map $r_v \cup [r_v \times 1_{\bigcup_i Z_i}]$. Let X be the one-point compactification of X_0, and let r_v denote the extension of h to X. If v is of finite order k, then $(r_v)^k = 1_X$ and so r_v is chain transitive on X. If v is of infinite order, then $\{\infty\} = \omega(r_v)(x) = \alpha(r_v)(x)$, and so r_v is chain transitive on X in this case as well. The group $H(X)$ is isomorphic to the discrete group G by $v \mapsto r_v^{-1}$. □

In these cases, the homeomorphism group is discrete. It is possible to obtain rather large nondiscrete groups. We first review some standard topology constructions.

A *pointed space* is a pair (X, x) consisting of a space with a chosen *base point* $x \in X$. We let $H(X, x)$ denote the closed subgroup of $H(X)$ consisting of those homeomorphisms which fix x. A space Y can be regarded as a pointed space with base point an isolated point not in Y. If (X_1, x_1) and (X_2, x_2) are pointed spaces and $f : X_1 \to X_2$ is a function, we use the notation $f : (X_1, x_1) \to (X_2, x_2)$ to mean that $f(x_1) = x_2$.

If A is a nonempty closed subset of a space X, then the space X/A with A *smashed to a point* is the quotient space of X with respect to the closed equivalence relation $1_X \cup A \times A$. Thus, the quotient map $q : X \to X/A$ is a homeomorphism between the open sets $X \setminus A$ and $X/A \setminus \{x_A\}$ with x_A the point which is the image of A.

Given two pointed spaces (X_1, x_1), (X_2, x_2), their *smash product* is

$$(X_1, x_1) \# (X_2, x_2) = (X_{12}, x_{12}),$$

a pointed space consisting of the product $X_1 \times X_2$ with the *wedge* $X_1 \times \{x_2\} \cup \{x_1\} \times X_2$ smashed to the point x_{12}. We can also define the smash product of a pointed space (X_1, x_1) and any space X_2 as $(X_1, x_1) \# X_2 = (X_{12}, x_{12})$, where X_{12} is the product $X_1 \times X_2$ with $\{x_1\} \times X_2$ smashed to the point x_{12}. Notice that in this case we can regard the space X_{12} as the one-point compactification of $(X_1 \setminus x_1) \times X_2$. The projections $\pi_1 : (X_1, x_1) \# X_2 \to (X_1, x_1)$ and $\pi_2 : X_1 \setminus \{x_1\} \times X_2 \to X_2$ are open and surjective.

We can define the smash product of two continuous functions once we fix base points from the domains. For $i = 1, 2$, let X_i and Y_i be spaces, $f_i : X_i \to Y_i$ a continuous function and $x_i \in X_i$ a base point. We set $y_i = f_i(x_i)$ to obtain a pointed space (Y_i, y_i) for $i = 1, 2$. Let $(X_{12}, x_{12}) = (X_1, x_1) \# (X_2, x_2)$ and $(Y_{12}, y_{12}) = (Y_1, y_1) \# (Y_2, y_2)$, and let $q : X_1 \times X_2 \to X_{12}$ and $r : Y_1 \times Y_2 \to Y_{12}$ be the quotient maps. We define a continuous function $f = (f_1, x_1) \# (f_2, x_2)$ from (X_{12}, x_{12}) to (Y_{12}, y_{12}) by the formula $f \circ q = r \circ (f_1 \times f_2)$. We can also define the smash product of two functions when only one of the domains has a base point. If $(V, v) = (X_1, x_1) \# X_2$ and $(W, w) = (Y_1, y_1) \# Y_2$ and if $s : X_1 \times X_2 \to V$ and $t : Y_1 \times Y_2 \to W$ are the quotient maps, then we define the continuous function $g = (f_1, x_1) \# f_2$ from (V, v) to (W, w) by $g \circ s = t \circ (f_1 \times f_2)$.

Example 4.2. We construct a space X so that

- the homeomorphism group $H(X)$ contains a nontrivial path-connected subgroup;
- every $f \in H(X)$ is chain transitive.

Let (Z, e) be a chaotic space Z, as above, with base point $e \in Z$ such that $Z \setminus \{e\}$ is connected. Let W be a connected, compact manifold (perhaps with boundary) of positive dimension k. Let $(X, e_X) = (Z, e) \# W$. If $(z, w) \in (Z \setminus \{e\}) \times W = X \setminus \{e_X\}$ and h is a homeomorphism from an open set containing (z, w) into X, then $h(z, w) = (z_1, w_1)$ with $z_1 \in Z \setminus \{e\}$ implies $z = z_1$. If not, then we can choose disk neighborhoods of w and w_1 each homeomorphic to I^k, and we can choose disjoint open neighborhoods U of z and U_1 of z_1 so that h induces a homeomorphism from $U \times I^k$ onto a subset of $U_1 \times I^k$, and this contradicts condition RIG. If $h(z, w) = e_X$, then h would map points (z_2, w_2) close to (z, w) in $X \setminus \{e_X\}$ to points (z_1, w_1) with z_1 in $X \setminus \{e_X\}$ close to e. This does not happen by the previous argument. Thus, it follows that $\pi_1 \circ h = \pi_1$, where π_1 is the projection to (Z, e). This implies that $H(X) = H(X, e_X)$, and π_1 maps every $h \in H(X)$ to 1_Z.

Since the homeomorphisms of X leave the preimages of π_1 invariant, it follows that every $h \in H(X)$ is of the form $h(z, w) = (z, q(z)(w))$ with $q : Z \setminus \{e\} \to H(W)$ a continuous map. Thus, the space of continuous maps $C(Z \setminus \{e\}, H(W))$ with the obvious group structure is isomorphic as a group with $H(X)$. Notice that we need not worry about behavior as z approaches e in Z since all of $\{e\} \times W$ is smashed to a point. Thus, the isomorphism is topological if we choose an increasing sequence $\{K_n\}$ of compacta in Z with union $Z \setminus \{e\}$ and define the metric $d(q_1, q_2) = \sup_n 2^{-n} d_H(q_1|_{K_n}, q_2|_{K_n})$ with d_H the uniform metric on $H(W)$.

In particular, the constant maps q yield $H(W)$ as a subgroup of $H(X)$. Since W is a manifold of positive dimension, it follows that the path components of the identity in $H(W)$ and hence in $H(X)$ are nontrivial subgroups. The remaining path components are cosets and so are nontrivial as well.

Let $h \in H(X)$. Given $\epsilon > 0$ and $(z, w) \in X \setminus \{e_X\}$, there is an ϵ-chain $z_0, \dots, z_N$ for 1_Z with $z_0 = z$ and $z_N = e$ and $z_i \neq e$ for $i < N$. That is, $d(z_{n+1}, z_n) \leq \epsilon$. Define $\{w_0, \dots, w_N\}$ by $w_0 = w$ and $h(z_i, w_i) = (z_i, w_{i+1})$ for $i < N$. Clearly, $(z_0, w_0), \dots, (z_{N-1}, w_{N-1}), e_X$ is an ϵ-chain for h from (z, w) to e_X. Similarly, there is an ϵ-chain for h^{-1} from (z, w) to e_X. It follows that h is chain recurrent. Since X is connected, h is chain transitive. □

Now let (Y, e) be a pointed space such that Y and $Y \setminus \{e\}$ are connected and, for every $y \in Y$, the open set $Y \setminus \{y\}$ has only finitely many components. Let C be a zero-dimensional space, and let $(X, e_X) = (Y, e) \# C$. Since C is zero-dimensional, the components of $X \setminus \{e_X\}$ are the sets $\{(Y \setminus \{e\}) \times \{c\} : c \in C\}$. If $y \in Y \setminus \{e\}$ and $c \in C$, then the components of $X \setminus \{(y, c)\}$ which do not contain e are all of the form $D \times \{c\}$, where D is a component of $Y \setminus \{y\}$ which does not contain e. Thus, if C is infinite, $X \setminus \{e_X\}$ has infinitely many components, while $X \setminus \{(y, c)\}$ has only finitely many components for $y \in Y \setminus \{e\}$, $c \in C$. Hence, if C is infinite, then any homeomorphism of X fixes e_X. We will assume that, even with C finite, the space Y is such that the point e_X is fixed by every homeomorphism of X.

It follows that for any homeomorphism on h on X, the projection $\pi_2 : X \setminus \{e_X\} \to C$ maps the restriction of h on $X \setminus \{e_X\}$ to a homeomorphism on C. The map on C is continuous because with $y_0 \neq e$ fixed, the map is given by $c \mapsto \pi_2(h(y_0, c))$. Similarly, the inverse is continuous. Thus, we obtain $(\pi_2)_* : H(X) \to H(C)$, a continuous, surjective homomorphism of topological groups. This splits via the continuous injection $j : H(C) \to H(X)$ given by $j(k) = (1_Y, e) \# k$.

Now suppose that $h \in H(X)$ is in the kernel of $(\pi_2)_*$, that is, it projects to 1_C. This means that every $(Y \setminus \{e\}) \times \{c\}$ is h-invariant. It follows that $h(y, c) = (q(c)(y), c)$, where $q : C \to H(Y, e)$ is a continuous map. That is, the kernel is $C(C, H(Y, e))$ with the obvious topological group structure. Thus, $H(X)$ is the semidirect product of $H(C)$ with $C(C, H(Y, e))$. The adjoint action of $j(H(C))$ is just the action $H(C) \times C(C, H(Y, e)) \to C(C, H(Y, e))$ given by $(k, q) \mapsto q \circ (k^{-1})$.

Example 4.3. We construct a space X so that

- the homeomorphism group $H(X)$ is isomorphic to the homeomorphism group of the Cantor set;
- every $f \in H(X)$ is chain transitive.

Let (Z, e) be a pointed chaotic space as before, and let $(X, e_X) = (Z, e) \# C$ with C a zero-dimensional space. We first check that even if C is finite, any homeomorphism h on X fixes e_X. If not, then there exist points $x_1, x_2 \in Z$, distinct from each other and distinct from e, and points $a_1, a_2 \in C$ such that $h(x_1, a_1) = (x_2, a_2)$. This implies that h induces a homeomorphism between sufficiently small neighborhoods U_1 of x_1 and U_2 of x_2. Choosing these as disjoint neighborhoods, we obtain a contradiction of condition RIG.

In this case $H(Z, e) = H(Z) = \{1_Z\}$. That is, the group $H(Z, e)$ is trivial, and so the group $C(C, H(Z, e))$ is trivial. This means that $(p_2)_* : H(X) \to H(C)$ and $j : H(C) \to H(X)$ are inverse isomorphisms, so every homeomorphism on X is mapped by π_1 to 1_Z. Just as in Example 4.2, it follows that every homeomorphism is chain transitive.

When C is a Cantor set, we obtain an example with homeomorphism group isomorphic to the homeomorphism group of the Cantor set. $\square$

Because of the rigidity of the connecting links, it is not true in these examples that $H(X)$ acts transitively on X. We can obtain examples which satisfy this additional condition by using the beautiful construction of Slovak spaces due to Downarowicz, Snoha, and Tywoniuk in [Downarowicz et al. 2017].

Let g be a totally transitive homeomorphism on a Cantor set W. That is, g^n is topologically transitive for all $n \in \mathbb{Z} \setminus \{0\}$. The construction begins with the *suspension* of g. That is, let $Y = W \times [0, 1]$ with $(x, 1)$ identified with $(g(x), 0)$ for all $x \in W$. On Y we define the real flow $\phi : \mathbb{R} \times Y \to Y$, the associated time-$t$ map $\phi^t : Y \to Y$, and the path map $\phi_x : \mathbb{R} \to Y$ for $t \in \mathbb{R}$, $x \in W$ by

$$\begin{aligned} \phi(t, (x, s)) = \phi^t(x, s) &= (g^{[t+s]}(x), \{t + s\}), \\ \phi_x(t) &= \phi^t(x, 0), \end{aligned} \tag{4-1}$$

where $[a]$ and $\{a\}$ are the integer part and fractional part, respectively, of the real number a. Identifying W with $W \times \{0\} \subset Y$, we see that g on W is identified with the time-one map ϕ^1 restricted to W.

Observe that the time-s map of the flow ϕ^s restricts to a homeomorphism from $\left[-\frac{1}{3}, \frac{1}{3}\right] \times W$ onto a neighborhood of $W \times \{s\}$ in Y for $s \in [0, 1]$. It follows that the path components of Y are exactly the $\mathbb{R}$-orbits of the flow. For $x \in W$ we let $\mathbb{R}x = \phi_x(\mathbb{R})$ denote the $\mathbb{R}$-orbit through $(x, 0)$.

In Y there are three types of path components:

Type 1: If x is a periodic point for g, then $\mathbb{R}x$ is a circle embedded in Y_0. This is a *circle type* path component.

Type 2: If x is recurrent for neither g nor g^{-1}, i.e., $x \notin \omega g(x) \cup \alpha g(x)$, then $\mathbb{R}x$ is an embedded copy of $\mathbb{R}$. That is, ϕ_x is a homeomorphism from $\mathbb{R}$ onto its image in Y. This is an *embedded* $\mathbb{R}$ *type* path component.

Type 3: If x is not periodic but $x \in \omega g(x) \cup \alpha g(x)$, then ϕ_x is a continuous injection which is not a homeomorphism onto its image. In fact, $\mathbb{R}x \subset Y_0$ is not locally connected. This is an *injected* $\mathbb{R}$ *type* path component.

The Slovak space construction is based on the following result:

Theorem 4.4 [Downarowicz et al. 2017, Lemma 4.3]. *Let f be a homeomorphism on a space Y with y_0 a point of Y which is not a periodic point for f. Let $\{a_n : n \in \mathbb{Z}\}$ be a sequence of positive reals such that $\sum_n a_n = 1$ and the set $\{|\ln(a_n) - \ln(a_{n-1})| : n \in \mathbb{Z}\}$ is bounded, and let $u : Y \setminus \{y_0\} \to [0, 1]$ be a continuous function. Let $Y' = Y \setminus \mathcal{O} f_{\pm}(y_0)$, and on Y' define the continuous function $u' = \sum_n a_n u \circ f^n$ so that the graph of u' is a closed subset of $Y' \times [0, 1]$ with the first coordinate projection $p : u' \to Y'$ a homeomorphism. Define on the set u' the homeomorphism $f' = (p)^{-1} \circ f \circ (p)$. Let X be the closure of u' in $Y \times [0, 1]$.*

The homeomorphism f' and its inverse are uniformly continuous on u', and so f' extends to a homeomorphism h on X. The first coordinate projection $p : X \to Y$ maps h on X to f on Y. If $y \in Y'$, then $(y, u'(y))$ is the unique point of X which is mapped by p to y. Hence, h is an almost one-to-one extension of f.

Now, following [Downarowicz et al. 2017], we apply this construction to the situation above, i.e., with Y the suspension of the Cantor set W via a totally transitive homeomorphism g.

Let x_0 be an element of the dense G_δ set $\bigcap\{\mathrm{Trans}(g^n) : n \in \mathbb{Z} \setminus \{0\}\}$. By [Akin 1993, Proposition 6.3(a)] the set of $\tau \in (0, \infty)$ such that $\omega\phi^{n\tau}(x_0, 0) = Y = \alpha\phi^{n\tau}(x_0, 0)$ is residual in $(0, \infty)$. Since the irrationals are also residual, we may fix such a τ, irrational in $(0, 1)$, and let $f = \phi^\tau$. Thus, $y_0 = (x_0, 0)$ is a transitive point for f^n on Y for all $n \in \mathbb{Z} \setminus \{0\}$. Since τ is irrational, f has no periodic points.

We first define u on a piece of the orbit of the point $y_0 = (x_0, 0)$:

$$u(\phi_{x_0}(t)) = \begin{cases} 0 & \text{for } -\frac{1}{2} \le t < 0, \\ \frac{1}{2}\big(1 - \cos\big(\frac{\pi}{t}\big)\big) & \text{for } 0 < t \le \frac{1}{2}. \end{cases} \tag{4-2}$$

Apply the Tietze extension theorem to obtain the continuous function $u : Y \setminus \{y_0\} \to [0, 1]$.

Apply Theorem 4.4 to Y with the homeomorphism f. We obtain an almost one-to-one lift h on $X = \overline{u'} \subset Y \times [0, 1]$. All of the path components of X are mapped by p homeomorphically onto the path components of Y, except that the Type 3 path component $\mathbb{R}y_0$ is cut into a sequence of path components of a new type.

Type 4: The path component Comp_n of $h^n(y_0)$ is mapped by p onto the set $\phi_{x_0}\big(((n-1)\tau, n\tau]\big)$. As $t \searrow (n-1)\tau$, above the open interval end, there is a topologist's sine which projects homeomorphically. Above the $n\tau$ endpoint there is a vertical segment $J_n = \{h^n(y_0)\} \times [0, a_{-n}]$ to which the oscillating end of the path component Comp_{n+1} converges. That is, $J_n = \mathrm{Comp}_n \cap \overline{\mathrm{Comp}_{n+1}}$. Thus, each path component Comp_n is a homeomorphic image of $\mathbb{R}_+ = [0, \infty)$. Each is an *embedded* $\mathbb{R}_+$ path component.

Theorem 4.5. *The homeomorphism group of X is $H(X) = \{h^n : n \in \mathbb{Z}\}$. For all $n \neq 0$ the homeomorphism h^n is topologically transitive.*

Proof. If $n \neq 0$, then by choice of τ, the homeomorphism $f^n = \phi^{n\tau}$ is topologically transitive on Y. Because p mapping h to f is an almost one-to-one lift, it follows that h^n is topologically transitive on X.

If h_1 is any homeomorphism on X, then the Type 4 component Comp_n is mapped to some Type 4 component Comp_{n+k}. Furthermore, $J_n = \mathrm{Comp}_n \cap \overline{\mathrm{Comp}_{n+1}}$ is mapped to J_{n+k}. It follows that p projects h_1 to a continuous map on Y which agrees with f^k on the dense set $\mathcal{O}(f_\pm)(y_0)$. It follows that it projects to f^k. Since p is almost one-to-one, it follows that $h_1 = h^k$. □

Downarowicz, Snoha, and Tywoniuk begin with g on W minimal and observe that, for a residual set of positive reals τ, the homeomorphism ϕ^τ is minimal on Y.

Choosing one such τ, they have f minimal on Y. All of the components of Y are then of Type 3. Then h^n is minimal on X for all $n \neq 0$.

We recall and extend their definition of a Slovak space.

Definition 4.6. (a) A space X is a *Slovak space* if X contains at least three points, $H(X)$ is isomorphic to $\mathbb{Z}$ and every $h \in H(X) \setminus \{1_X\}$ is minimal.

(b) A space X is *Slovakian* if X contains at least three points, $H(X)$ is nontrivial and every $h \in H(X) \setminus \{1_X\}$ is topologically transitive.

We extend [Downarowicz et al. 2017, Theorem 4] with essentially the same proof.

Theorem 4.7. *A Slovakian space is connected, and its homeomorphism group has no elements of finite order other than the identity.*

Proof. Let h be a topologically transitive homeomorphism on a space X. Suppose that X contains a proper, clopen, nonempty subset A. If $h^{-1}(A) \subset A$, then for any $x \in \mathrm{Trans}(h)$, we have $h^n(x) \in A$ for some $n \in \mathbb{N}$, and so $x \in A$. Thus, the dense set $\mathrm{Trans}(h)$ is contained in A, contradicting the assumption that A is a proper, clopen set. It follows that $B = h^{-1}(A) \setminus A$ is a proper, clopen, nonempty set, and $B \cap h(B) = \varnothing$. Define

$$g(x) = \begin{cases} h(x) & \text{for } x \in B, \\ h^{-1}(x) & \text{for } x \in h(B), \\ x & \text{for } x \in X \setminus (B \cup h(B)). \end{cases} \tag{4-3}$$

The points of the nonempty set $B \cup h(B)$ are periodic with period 2, and so $g \neq 1_X$. Since $g^2 = 1_X$, it is clear that g is not topologically transitive.

Since X is nontrivial and connected, it is perfect and therefore uncountable. On such a space, no topologically transitive homeomorphism has finite order. □

Questions. Does there exists a Slovakian space X for which $H(X)$ is not discrete? More generally, does there exist a nontrivial space X such that the topologically transitive homeomorphisms are dense in $H(X)$? If X is such a space, then every $h \in H(X)$ is chain transitive. In particular, since 1_X is chain transitive, X is connected. On the other hand, 1_X is not topologically transitive, but it is a limit of topologically transitive homeomorphisms, and so $H(X)$ is not discrete.

The only Slovakian spaces we know of are variations on the original construction of [Downarowicz et al. 2017]. All of these have homeomorphism group isomorphic to $\mathbb{Z}$.

Now we extend the above construction, which was built on a totally transitive homeomorphism g on a Cantor space W. Suppose that we are given B, a proper, closed, g-invariant subset of W, and that $r : B \to A$ is a continuous surjection which maps the restriction $g|_B$ to 1_A, the identity on the space A. That is, $r^{-1}(a)$ is a closed, invariant set in B for every $a \in A$. Since B is proper, closed and invariant, it is disjoint from $\mathrm{Trans}(g)$. In particular, $x_0 \notin B$.

Let $\widehat{B}$ be the quotient of $B \times [0, 1]$ in Y. This is just the suspension of $g|_B$. Since $r \circ g = r$ on B, it follows that $r \circ \pi_1 : B \times [0, 1] \to A$ factors to define the surjection $\hat{r} : \widehat{B} \to A$, and each $\hat{r}^{-1}(a)$ is a ϕ-invariant closed subset of Y. Since $x_0 \notin B$ and B is g-invariant, $\widehat{B} \subset Y'$.

The preimage of $\widehat{B}$ via the homeomorphism $p : u' \to Y'$ is a compact f'-invariant subset of $X = \overline{u'}$. Thus, $p^{-1}(\widehat{B})$ is a closed h-invariant subset of X.

Now we attach Y and X to A, using the maps $\hat{r}$ and $\hat{r} \circ p$. That is, we define

$$\begin{aligned} E_{r,Y} &= 1_Y \cup [(\hat{r})^{-1} \circ \hat{r}], \\ E_{r,X} &= 1_X \cup [(\hat{r} \circ p)^{-1} \circ (\hat{r} \circ p)], \end{aligned} \tag{4-4}$$

closed equivalence relations on Y and X, respectively. Let $q_r^Y : Y \to Y_r$ and $q_r^X : X \to X_r$ be the projections to the quotient spaces. The homeomorphisms f and h induce homeomorphisms f_r and h_r on the quotient spaces. The projection $p : X \to Y$ induces $p_r : X_r \to Y_r$, a continuous, almost one-to-one surjection which maps h_r to f_r. As usual, we will regard the homeomorphisms induced by $\hat{r} : \widehat{B} \to A$ and $\hat{r} \circ p : p^{-1}(\widehat{B}) \to A$ as identifications, so A is thought of as a subset of Y_r and also as a subset of X_r. Recall that f has no periodic points and so h does not either. It follows that $A \subset Y$ and $A \subset X$ are the sets of fixed points for f_r and h_r, respectively.

For $a \in A$ and $x \in W$, we say that $q_r^Y(\mathbb{R}x)$ is an *a-orbit* if $\omega g(x) \subset r^{-1}(a)$ or $\alpha g(x) \subset r^{-1}(a)$. If $\omega g(x) \subset r^{-1}(a)$, then the map $q_r^Y \circ \phi_x : \mathbb{R} \to Y_r$ extends continuously to $\mathbb{R} \cup \{+\infty\}$ by mapping $+\infty$ to a. If $x \in B$, then x is an a-orbit if and only if $r(x) = a$, in which case $q_r^Y(\mathbb{R}x) = \{a\}$. If K_1 and K_2 are path components of A, they are *linked* if there exists an $x \in W$ which is both an a_1-orbit and an a_2-orbit for some $a_1 \in K_1$, $a_2 \in K_2$, i.e., if $\alpha g(x) \subset r^{-1}(a_1)$ and $\omega g(x) \subset r^{-1}(a_2)$ or vice versa. Two path components are *linkage equivalent* if there is a finite sequence $K_1, \ldots, K_N$ joining them with each K_i linked to its successor.

In Y_r we have a new path component type:

Type 5: Let $[K]$ be a linkage equivalence class of path components of A. The $[K]$-*component* in Y is the union of all of the a-orbits for $a \in K_1 \in [K]$ and of the path components $K_1 \in [K]$.

It is possible for a $[K]$-component to be of embedded $\mathbb{R}_+$ type. The only way this can happen is if the embedding of $[0, \infty)$ onto the $[K]$-component maps 0 to a point of A. The *endpoint* of this $\mathbb{R}_+$ type component lies in A.

Theorem 4.8. *The homeomorphism group of X_r is $H(X_r) = \{h_r^n : n \in \mathbb{Z}\}$. The homeomorphism h_r^n is topologically transitive for all $n \neq 0$.*

Proof. The surjection q_r^X maps the totally transitive homeomorphism h onto the totally transitive homeomorphism h_r.

If h_1 is any homeomorphism on X_r, we proceed just as in the proof of Theorem 4.5. There is, however, one tricky bit. It is possible that a Type 5 path component is of embedded $\mathbb{R}_+$ type. If Comp_n maps to Comp_{n+k}, then just as before, $J_n = \mathrm{Comp}_n \cap \overline{\mathrm{Comp}_{n+1}}$ is mapped to J_{n+k}, and $h_1 = h^k$ as before.

Suppose instead that Comp_n is mapped to some Type 5 path component, which we call Q_n. Then Comp_{n+1} will have to be mapped to a Type 5 path component Q_{n+1} with J_n mapping to $Q_n \cap \overline{Q_{n+1}}$, and this set contains the endpoint of Q_n which is in A. Thus, there is a sequence $\{x_n \in J_n : n \in \mathbb{Z}\}$ with $h_1(x_n) \in A$. But the sequence $\{x_n\}$ projects to the f_r-orbit of y_0, and this is dense in Y_r. Since $p_r : X_r \to Y_r$ is an almost one-to-one map, the sequence $\{x_n\}$ is dense in X_r. Since the proper closed subset A contains the sequence $\{h_1(x_n)\}$ and h_1 is a homeomorphism, we obtain a contradiction. □

If X is one of the examples of Slovakian spaces as constructed above, then $X \setminus D$ is connected for any countable subset D of X. This is because we can choose a countable invariant subset $D_0 \subset W$ with $x_0 \in D_0$ such that $D \subset D_0 \times [0, 1]$. Since $\mathrm{Trans}(g) \cap \mathrm{Trans}(g^{-1})$ is residual and thus uncountable, we can choose $x \in (\mathrm{Trans}(g) \cap \mathrm{Trans}(g^{-1})) \setminus D_0$. Then the orbit $\mathbb{R}x$ is connected, dense in X, and contained in $X \setminus D$. We don't know whether this property holds for all Slovakian spaces.

Now for our final construction.

Example 4.9. Let C be a space which is countable or the Cantor set. We construct a space K so that

- the homeomorphism group $H(K)$ is isomorphic as a topological group to the semidirect product of $H(C)$ with $C(C, \mathbb{Z})$;
- the action of $H(K)$ on K is topologically transitive;
- every element of $H(K)$ is chain transitive;
- if C is either finite or the Cantor set, then there exists a topologically transitive homeomorphism in $H(K)$.

Let g be a totally transitive homeomorphism on a Cantor set W with a fixed point $e \in W$. Let $r : \{e\} \to \{e\}$ be the identity. The space Y_r is the suspension of g with the circle $\{e\} \times [0, 1]$ smashed to the point e, and ϕ_r is the associated real flow on Y_r. Then X_r is the Slovakian space over Y_r with $p_r : X_r \to Y_r$ the almost one-to-one projection. The group $H(X_r)$ is cyclic with generator h_r mapped by p to ϕ_r^τ.

Let (K, e_K) be the smash product $(X_r, e) \# C$, and let $(L, e_L) = (Y_r, e) \# C$. We have the projections

$$\begin{aligned} \pi_L &: L \setminus \{e_L\} = (Y_r \setminus \{e\}) \times C \to C, \\ \pi = \pi_L \circ (p \times 1_C) &: K \setminus \{e_K\} = (X_r \setminus \{e\}) \times C \to C, \end{aligned} \tag{4-5}$$

and $P : (K, e_K) \to (L, e_L)$ is the smash product $(p, e) \# 1_C$.

If C is a singleton, then (K, e_K) is just (X_r, e), $H(C)$ is trivial, and $C(C, \mathbb{Z})$ is isomorphic to $\mathbb{Z}$ and to $H(K)$. The identity is chain transitive, and all other elements of $H(K)$ are topologically transitive. So the result is clear in this case. Now assume that C has at least two points. This implies that e_K disconnects K, while no other point does since no point of X_r disconnects X_r. Hence, every homeomorphism of K preserves e_K.

From the discussion preceding Example 4.3 it follows that $H(K)$ is the semidirect product of $H(C)$ and $C(C, H(X_r))$, which is essentially $C(C, \mathbb{Z})$ since X_r is Slovakian. The projection $\pi : K \setminus \{e_K\} \to C$ induces the group surjection $\pi_* : H(K) \to H(C)$, which is split by the injection $j : H(C) \to H(K)$. If $k \in H(C)$, then $j(k)$ is $(1_{X_r}, e) \# k$. In this case, the subgroup $C(C, \mathbb{Z})$ is commutative. If $q \in C(C, \mathbb{Z})$, then the associated element of $H(K)$ is the projection of $(x, c) \mapsto (h_r^{q(c)}(x), c)$. The constant elements, the homeomorphisms $(h^n, e) \# 1_C$, commute with all the elements of $H(K)$ since they commute with the members of the subgroup $j(H(C))$. For $(n, k) \in \mathbb{Z} \times H(C)$ let $J(n, k) = (h^n, e) \# k$. Thus, $J : \mathbb{Z} \times H(C) \to H(K)$ is a topological embedding and a group homomorphism.

Note that P maps $(h_r^{q(c)}(x), c)$ to $(\phi_r^{q(c)\tau}(x), c)$ and maps $(y, k(c))$ to $(p(y), k(c))$. It follows that every homeomorphism of K projects by P to a homeomorphism of L.

Case 1 (C is finite): Let k be a cyclic permutation on C so that C consists of a single periodic orbit under k. Since h is totally transitive, the product $h \times k$ on $X_r \times C$ is topologically transitive, and so it projects to a topologically transitive element of $\mathbb{Z} \times H(C) \subset H(K)$.

Let F be an arbitrary homeomorphism on K with $k = \pi_*(F)$. Since k is a permutation of a finite set, there exists $N \in \mathbb{N}$ such that k^N is the identity. This means that F^N projects to the identity on C and so preserves each fiber $X_r \times \{c\}$. So there exists $n_c \in \mathbb{Z}$ such that F^N restricts to h^{n_c} on the fiber. This is chain transitive on the fiber, topologically transitive if $n_c \neq 0$, and so every point is chain recurrent for F^n and hence for F. Since K is connected, F is chain transitive on K.

Case 2 (C is countably infinite): Since C is countable, the isolated points are dense in C. If $x_1, x_2 \in C$ are distinct isolated points, let k interchange these two points and fix the remaining points of C. By Case 1, $J(1, k) = (h, e) \# k$ restricts to a topologically transitive homeomorphism on the closed subset $(X_r, e) \# \{x_1, x_2\}$. This implies that $H(K)$ acts in a topologically transitive manner on the set $(X_r \setminus \{e\}) \times \mathrm{Iso}(C)$. This is a dense open subset of K, and so $H(K)$ is topologically transitive on K.

Case 3 (C is a Cantor set): A homeomorphism k on C is *topologically mixing* if, for all nonempty open sets $U_1, U_2 \subset C$, there exists $N \in \mathbb{N}$ such that $N_k(U_1, U_2) = \{n : k^n(U_1) \cap U_2 \neq \varnothing\}$ contains every $m \geq N$. The shift homeomorphism on $\{0, 1\}^{\mathbb{Z}}$ is topologically mixing, and the product of a topologically mixing and a topologically transitive homeomorphism is topologically transitive. It follows that if $k \in H(C)$ is topologically mixing, then $J(1, k)$ is a topologically transitive element of $H(K)$.

It remains to show that every element of $H(K)$ is chain transitive. It suffices to show that, for an arbitrary $F \in H(K)$ and an arbitrary point y of K, we have $e_K \in \mathcal{C}F(y)$. Applying this to F and F^{-1}, we see that every point of K is chain recurrent for F.

With the homeomorphism g on W chosen arbitrarily subject to the above conditions, we do not know whether this is always true. We prove it by imposing further restrictions on the homeomorphism g with which we began.

Call a homeomorphism g *semiminimal* if it has a fixed point e and if for every $x \neq e$ the orbit $\mathcal{O}g_{\pm}(x) = \{g^n(x) : n \in \mathbb{Z}\}$ is dense in W. Such semiminimal homeomorphisms exist. Topologically mixing examples of semiminimal homeomorphisms on the Cantor set are constructed explicitly in [Akin 2016, Theorem 4.19], and Theorem 4.16 of that paper shows that the semiminimal homeomorphisms form a dense G_δ subset of the set of those chain transitive homeomorphisms on a Cantor set which admit a fixed point. Now assume that g is a semiminimal homeomorphism which is topologically mixing and so is totally transitive. This implies that if $x \neq e$, then the real orbit $\mathbb{R}x$ is dense in Y_r.

Let $F \in H(K)$ and $y \in K$. We must show that y chains to e_K. Let $k = \pi_*(F)$ be the induced homeomorphism on C and $G = P_*(F)$ the induced homeomorphism on L.

The invariant set $\omega F(y)$ contains a closed subset M such that the restriction of F to M is minimal. Of course, $M \subset \mathcal{C}F(y)$. So it suffices to prove that $e_K \in \mathcal{C}F(M)$. This is obvious if $M = \{e_K\}$. Now assume M is not equal to $\{e_K\}$ and so does not contain e_K since distinct minimal sets are disjoint.

Hence, M is a compact subset of $K \setminus \{e_K\}$. Since π maps F to k, the subset $Q = \pi(M)$ of C is compact and invariant on which k restricts to a minimal homeomorphism. Let $(\widehat{K}, e_K) = (X_r, e)\#Q$ and $(\widehat{L}, e_K) = (Y_r, e)\#Q$, so that $\widehat{K}$ is a closed F-invariant subset of K and $\widehat{L}$ is a closed G-invariant subset of L. The restriction of F to $M \subset \widehat{K}$ is minimal, and so if $\widetilde{M} = P(M) \subset \widehat{L}$, then G on $\widetilde{M}$ is minimal. On $\widetilde{L}$ (but not on $\widetilde{K}$) the homeomorphism $(\phi_r^t, e)\,\#\,1_Q$ is defined for every t, and each such homeomorphism commutes with G. It follows that $\widetilde{M}_t = ((\phi_r^t, e)\#1_Q)(\widetilde{M})$ is a G-invariant subset on which G is minimal. If $a \in A$, there exists $((x, s), a) \in \widetilde{M}$ for some $x \in W \setminus \{e\}$ and $s \in [0, 1)$ because π_L maps $\widetilde{M}$ onto A. Because $x \neq e$, the real orbit $\mathbb{R}x$ is dense in Y_r. It follows that $\bigcup_t \widetilde{M}_t$ is dense in every fiber $\pi_L^{-1}(a)$. This implies that the union of the minimal subsets of G is dense in $\widehat{L}$. Above each $\widetilde{M}_t$ there is a minimal subset $M_t \subset \widehat{K}$ with $P(M_t) = \widetilde{M}_t$. Because P is an almost one-to-one map, it follows that the union $\bigcup_t M_t$ is dense in $\widehat{K}$. This implies that the recurrent points for F are dense in $\widehat{K}$, and so every point of $\widehat{K}$ is chain recurrent. Since $\widehat{K}$ is connected, F on $\widehat{K}$ is chain transitive and, in particular, $e_K \in \mathcal{C}F(M)$, as required. □

Remark. Notice that an almost one-to-one lift of a chain transitive map need not be chain transitive. Let t be translation by 1 on $\mathbb{Z}$ and t^*, t^{**} the extensions to the one-point compactification $\mathbb{Z}^* = \mathbb{Z} \cup \{\infty\}$ and the two-point compactification $\mathbb{Z}^{**} = \mathbb{Z} \cup \{+\infty, -\infty\}$. The map $p : \mathbb{Z}^{**} \to \mathbb{Z}^*$ which maps both $+\infty$ and $-\infty$ to ∞ is an almost one-to-one map, but while t^* is chain transitive, t^{**} is not.

In general, if A is a nowhere dense, closed, invariant set which contains $|\mathcal{C}f|$ for a homeomorphism f on X, then smashing A to the point e in X/A, the induced homeomorphism f_A on X/A has e as the unique minimal point, so f_A is chain transitive, the projection $q_A : X \to X/A$ is almost one-to-one since A is nowhere dense, and f is not chain recurrent since $|\mathcal{C}f|$ is a proper subset of X. For example, with $X = I^2$ and $f(x, y) = (x^2, y^2)$, the corner points are the only chain recurrent points. Let A be the boundary $I \times \{0, 1\} \cup \{0, 1\} \times I$.

Appendix: Chaotic spaces

The rigid and strongly chaotic spaces constructed in [de Groot and Wille 1958] and [Charatonik and Charatonik 1996] are dendrites or subsets of $\mathbb{R}^2$. It will be convenient for our purposes to use infinite-dimensional examples for the sequence of spaces $\{Z_n\}$ which satisfy RIG and CON.

Let M be an infinite subset of $\mathbb{N} \setminus \{1\}$, and let $m : \mathbb{N} \to M$ be the unique order-preserving bijection. For $n = 0, 1, \ldots,$ let $M^n = \{m(2^n(2k-1)) : k \in \mathbb{N}\}$ so that $\{M^n\}$ is a partition of M by a pairwise disjoint sequence of infinite sets. For all $i \in \mathbb{N}$ let S^i be the sphere in $\mathbb{R}^{i+1}$ of radius i^{-1} centered at $(i^{-1}, 0, \ldots, 0)$ so that the origin 0 is a point of S^i. Let $S(n, k) = S^{m(2^n(2k-1))}$ for $n = 0, 1, \ldots$ and $k \in \mathbb{N}$.

Let Z^0 be the two-torus $S^1 \times S^1$, and let $A^0 = \{a^0_k : k \in \mathbb{N}\}$ be a dense sequence of distinct points in Z^0. Let Z^1 be Z^0 with a copy of $S(0, k)$ attached to Z^0 with $0 \in S(0, k)$ identified with the *attachment point* a^0_k for each $k \in \mathbb{N}$ so that the attached spheres are disjoint in Z^1. Let $r^1 : Z^1 \to Z^0$ be the retraction with each $S(0, k)$ mapped to the point a_k. For each $k \in \mathbb{N}$ let $r(0, k) : Z^1 \to S(0, k)$ be the retraction mapping $Z^1 \setminus S(0, k)$ to the point a_k. Since the diameters of the spheres tend to zero, the space Z^1 is compact and metrizable and the retractions are continuous. The open set $Z^1 \setminus Z^0$ is dense in Z^1.

For $n \geq 1$ assume that Z^n has been defined with a retraction $r^n : Z^n \to Z^{n-1}$ and with retractions $r(n-1, k) : Z^n \to S(n-1, k)$ onto the attached spheres and so that $Z^n \setminus Z^{n-1}$ is dense in Z^n. Let $A^n = \{a^n_k : k \in \mathbb{N}\}$ be a sequence of distinct points dense in $Z^n \setminus Z^{n-1}$. Let Z^{n+1} be Z^n with a copy of $S(n, k)$ attached to Z^n with $0 \in S(n, k)$ identified with the attachment point a^n_k for each $k \in \mathbb{N}$ so that the attached spheres are disjoint in Z^{n+1}. Let $r^{n+1} : Z^{n+1} \to Z^n$ be the retraction which maps the new spheres to their attachment points, and let $r(n,k) : Z^{n+1} \to S(n,k)$ be the retraction mapping $Z^{n+1} \setminus S(n,k)$ to a^n_k. Notice that for each $x \in Z^n \setminus A^n$ the set $(r^{n+1})^{-1}(x)$ equals $\{x\}$.

Let Z be the inverse limit of the system $\{r^n : Z^n \to Z^{n-1}, n \in \mathbb{N}\}$. The inclusions $i_n : Z^n \to Z^m$ with $m > n$ commute with the retractions. We obtain a limiting inclusion and so can regard $\{Z^n\}$ as an increasing sequence of subsets of Z. The projection $r_n : Z \to Z^n$ is then a retraction with $r^n \circ r_n = r_{n-1}$. We write $r(n,k) : Z \to S(n,k)$ for the composition of retractions $r(n,k) \circ r_{n+1}$. Notice that if we pick z from the set $Z^* = Z \setminus \big(\bigcup_n Z^n\big)$, then there is a unique sequence of attachment points $\{a^n_{k_n}\}$ converging to z with $r^n(a^n_{k_n}) = a^{n-1}_{k_{n-1}}$ for all n. Thus, Z^*, while a dense G_δ, is totally disconnected.

Observe that $(r_{n+1})^{-1}(S(n,k))$ and $\{a^n_k\} \cup (r_n)^{-1}(Z^n \setminus \{a^n_k\})$ are connected sets with union Z and which meet only at a^n_k. Thus, each attachment point disconnects Z.

If $F \subset Z$ is a finite set containing no attachment points, then $Z \setminus F$ is connected. For any finite $F \subset Z$, the set $Z \setminus F$ contains only finitely many components. In fact, the number of components is exactly one more than the number of attachment points in F. Thus, we obtain the condition CON.

Lemma A.1. *If W is a closed, connected, nontrivial subset of $Z \times I^N$ such that $W \setminus A$ is connected for any $A \subset W$ with topological dimension at most N, then either $W \subset Z^0 \times I^N$ or there exists a unique attached sphere $S(n,k)$ such that $W \subset S(n,k) \times I^N$.*

Proof. Notice first that the dimension of W is at least $N+2$. For if U is a nonempty open set with $\overline{U}$ a proper subset of W, then the topological boundary of U disconnects W and so has dimension at least $N+1$. This implies that the dimension of W is at least $N+2$.

The set W is not contained in $Z^* \times I^N$ since Z^* is totally disconnected and the components of $Z^* \times I^N$ have dimension N. Assume W is not a subset of $Z^0 \times I^N$.

If W meets $(Z^{n+1} \setminus Z^n) \times I^N$ for some $n \geq 0$, then X meets $(S(n,k) \setminus \{a^n_k\}) \times I^N$ for some k. Since $a^n_k \times I^N$ disconnects $Z \times I^N$ and no such set disconnects W, it follows that $W \subset r_n^{-1}(S(n,k)) \times I^N$. For the attachment points a^{n+1}_j in $S(n,k) \setminus \{a^n_k\}$, the sets $a^{n+1}_j \times I^N$ also disconnect Z. We see that either $W \subset S(n,k) \times I^N$ or else $W \subset r_{n+1}^{-1}(b) \times I^N$, where b is one of the attachment points in $S(n,k) \setminus \{a^n_k\}$.

If this process does not halt with W contained in the product of I^N with some attached sphere, then there is a sequence of attachment points $\{b^{n_i}_{k_i}\}$ with $W \subset r_{n_i+1}^{-1}(b^{n_i}_{k_i}) \times I^N$ and $r_{n_i+1}(b^{n_{i+1}}_{k_{i+1}}) = b^{n_i}_{k_i}$ with $n_i \to \infty$. The limit of such a sequence $\{b^{n_i}_{k_i}\}$ is a point of Z^*, and this would imply that W is a subset of $Z^* \times I^N$. $\square$

If A and X are spaces, a *nonnull embedding* of A in X is a continuous injective function $j : A \to X$ which is not homotopic to a constant map in X.

Corollary A.2. *If S is a sphere of dimension at least two and $j : S \times I^N \to Z \times I^N$ is a nonnull embedding, then for a unique pair (n, k), we have $j(S \times I^N) \subset S(n, k) \times I^N$. Furthermore, the dimension of S equals the dimension of $S(n, k)$. On the other hand, if $j_0 : S \to S(n, k)$ is a homeomorphism, then $j = j_0 \times 1_{I^N} : S \times I^N \to Z \times I^N$ is a nonnull embedding.*

Proof. Corollary 1 of Theorem IV.4 in [Hurewicz and Wallman 1941] says that a connected manifold of dimension at least $N + 2$ cannot be disconnected by a subset of dimension N. Hence, Lemma A.1 applies to $W = j(S \times I^N)$, and so $j(S \times I^N) \subset S(n, k) \times I^N$ for some pair (n, k) which is clearly unique. The space $S \times I^N$ cannot be embedded in a manifold of smaller dimension, and so $\dim S \le \dim S(n, k)$. If the inequality were strict then the map j would be homotopically trivial since the homotopy groups of a sphere vanish below its dimension. Hence, the dimension of S must be $m(2^n(2k-1)) = \dim S(n, k)$.

If $j_0 : S \to S(n, k)$ is a homeomorphism, then $j = j_0 \times 1_{I^N} : S \times I^N \to S(n, k) \times I^N$ is not homotopically trivial. Since $S(n, k)$ is a retract of Z, the embedding $j : S \times I^N \to Z \times I^N$ is not homotopically trivial. □

Thus, we can associate to any open subset $U \subset Z$ the set $\delta(U) = \{\dim S(n, k) : S(n, k) \subset U\} \subset M$. Since the diameters of the attaching spheres tend to zero, it follows that if U is nonempty, then $\delta(U)$ is infinite.

Corollary A.2 says that for a sphere S of dimension at least two there exists a nonnull embedding of $S \times I^N$ into $U \times I^N$ if and only if $\dim S \in \delta(U)$. That is, $\delta(U)$ is a topological invariant for the sets $U \times I^N$. Observe that if U_1 and U_2 are disjoint, nonempty, open sets in Z, then $\delta(U_1) \cap \delta(U_2) = \varnothing$ since distinct attached spaces have distinct dimensions.

If we begin by partitioning $\mathbb{N} \setminus \{1\}$ into a pairwise disjoint sequence $\{M_n\}$ of infinite subsets, then we can do this construction associating Z_n with M_n. That is, $\delta(Z_n) = M_n$. It follows that if $U_1 \subset Z_{n_1}$ and $U_2 \subset Z_{n_2}$ are nonempty open subsets with $n_1 \ne n_2$, then $\delta(U_1) \cap \delta(U_2) = \varnothing$.

Condition RIG of Section 4 is thus proved for this sequence of spaces.

Acknowledgements

We would like to express our appreciation to the referee for the helpful discussion of the history of the omega limit set problems and for suggestions concerning the exposition.

References

[Agronsky and Ceder 1991/92a] S. Agronsky and J. Ceder, “What sets can be ω-limit sets in E^n?”, *Real Anal. Exchange* **17**:1 (1991/92), 97–109. MR Zbl

[Agronsky and Ceder 1991/92b] S. Agronsky and J. G. Ceder, "Each Peano subspace of E^k is an ω-limit set", *Real Anal. Exchange* **17**:1 (1991/92), 371–378. MR Zbl

[Agronsky et al. 1989/90] S. J. Agronsky, A. M. Bruckner, J. G. Ceder, and T. L. Pearson, "The structure of ω-limit sets for continuous functions", *Real Anal. Exchange* **15**:2 (1989/90), 483–510. MR Zbl

[Akin 1993] E. Akin, *The general topology of dynamical systems*, Graduate Studies in Mathematics **1**, American Mathematical Society, Providence, RI, 1993. MR Zbl

[Akin 2016] E. Akin, "Conjugacy in the Cantor set automorphism group", pp. 1–42 in *Ergodic theory, dynamical systems, and the continuing influence of John C. Oxtoby*, edited by J. Auslander et al., Contemp. Math. **678**, American Mathematical Society, Providence, RI, 2016. MR Zbl

[Akin and Carlson 2012] E. Akin and J. D. Carlson, "Conceptions of topological transitivity", *Topology Appl.* **159**:12 (2012), 2815–2830. MR Zbl

[Andronov and Khaikin 1937] A. A. Andronov and S. E. Khaikin, Теория колебаний, Ob. Nauch.-Tekh. Izd. NKTP SSSR, Moscow and Leningrad, 1937. Translated as *Theory of oscillations*, Princeton Univ. Press, 1949. An enlarged second edition was published in 1959 by Fitmatgiz, Moscow, with A. A. Witt as a coauthor; translation: Pergamon Press, Oxford, 1966.

[Birkhoff 1927] G. D. Birkhoff, *Dynamical systems*, American Mathematical Society Colloquium Publications **9**, American Mathematical Society, Providence, RI, 1927. MR JFM

[Bruckner and Smítal 1992] A. M. Bruckner and J. Smítal, "The structure of ω-limit sets for continuous maps of the interval", *Math. Bohem.* **117**:1 (1992), 42–47. MR Zbl

[Charatonik and Charatonik 1996] J. J. Charatonik and W. J. Charatonik, "Strongly chaotic dendrites", *Colloq. Math.* **70**:2 (1996), 181–190. MR Zbl

[Conley 1978] C. Conley, *Isolated invariant sets and the Morse index*, CBMS Regional Conference Series in Mathematics **38**, American Mathematical Society, Providence, RI, 1978. MR Zbl

[Dowker 1953] Y. N. Dowker, "The mean and transitive points of homeomorphisms", *Ann. of Math.* (2) **58** (1953), 123–133. MR Zbl

[Dowker and Friedlander 1954] Y. N. Dowker and F. G. Friedlander, "On limit sets in dynamical systems", *Proc. London Math. Soc.* (3) **4** (1954), 168–176. MR Zbl

[Downarowicz et al. 2017] T. Downarowicz, L. Snoha, and D. Tywoniuk, "Minimal spaces with cyclic group of homeomorphisms", *J. Dynam. Differential Equations* **29**:1 (2017), 243–257. MR

[de Groot 1959] J. de Groot, "Groups represented by homeomorphism groups, I", *Math. Ann.* **138** (1959), 80–102. MR Zbl

[de Groot and Wille 1958] J. de Groot and R. J. Wille, "Rigid continua and topological group-pictures", *Arch. Math.* **9**:5 (1958), 441–446. MR Zbl

[Gutek 1979] A. Gutek, "On extending homeomorphisms on the Cantor set", pp. 105–116 in *Topological structures, II, Part I* (Amsterdam, 1978), edited by P. C. Baayen and J. van Mill, Math. Centre Tracts **115**, Math. Centrum, Amsterdam, 1979. MR Zbl

[Hurewicz and Wallman 1941] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton Mathematical Series **4**, Princeton Univ. Press, 1941. MR Zbl

[Knaster and Reichbach 1953] B. Knaster and M. Reichbach, "Notion d'homogénéité et prolongements des homéomorphies", *Fund. Math.* **40** (1953), 180–193. MR Zbl

[Kolyada and Snoha 1992/93] S. F. Kolyada and L. Snoha, "On ω-limit sets of triangular maps", *Real Anal. Exchange* **18**:1 (1992/93), 115–130. MR Zbl

[Lorch 1981] E. R. Lorch, "On some properties of the metric subalgebras of l^∞", *Integral Equations Operator Theory* **4**:3 (1981), 422–434. MR Zbl

[Lorch 1982] E. R. Lorch, "Certain compact spaces and their homeomorphism groups", *Rend. Sem. Mat. Fis. Milano* **52**:1 (1982), 75–86. MR Zbl

[Sharkovsky 1965] O. M. Sharkovskiĭ, "On attracting and attracted sets", *Dokl. Akad. Nauk SSSR* **160** (1965), 1036–1038. In Russian; translated in *Soviet Math. Dokl.* **6** (1965), 268–270. MR Zbl

[Sharkovsky et al. 1989] A. N. Sharkovskiĭ, S. F. Kolyada, A. G. Sivak, and V. V. Fedorenko, Динамика одномерных отображений, Naukova Dumka, Kiev, 1989. Translated as *Dynamics of one-dimensional maps*, Kluwer, Dordrecht, 1997. MR Zbl

[Tsankov 2006] T. Tsankov, "Compactifications of $\mathbb{N}$ and Polishable subgroups of S_∞", *Fund. Math.* **189**:3 (2006), 269–284. MR Zbl

Received December 2, 2015. Revised February 3, 2017.

ETHAN AKIN
MATHEMATICS DEPARTMENT
THE CITY COLLEGE
NEW YORK, NY
UNITED STATES
ethanakin@earthlink.net

JUHO RAUTIO
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF OULU
FINLAND
juhok.rautio@mail.suomi.net

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.51

SPINORIAL REPRESENTATION OF SUBMANIFOLDS IN RIEMANNIAN SPACE FORMS

Pierre Bayard, Marie-Amélie Lawn and Julien Roth

We give a spinorial representation of submanifolds of any dimension and codimension into Riemannian space forms in terms of the existence of generalized Killing spinors. We discuss several applications, among them a new and concise proof of the fundamental theorem of submanifold theory. We also recover results of T. Friedrich, B. Morel and the authors in dimensions 2 and 3.

1. Introduction

One of the fundamental problems in submanifold theory deals with the existence of isometric immersions from a Riemannian manifold M^n into another fixed Riemannian manifold N^{n+p}. If the ambient manifold is the space form $\mathbb{R}^{n+p}$, $\mathbb{S}^{n+p}$ or $\mathbb{H}^{n+p}$, the fundamental theorem of submanifold theory states that the Gauss, Ricci and Codazzi equations, also called structure equations, are necessary and sufficient conditions.

In the case of surfaces, another approach is given by the study of Weierstrass representations. Historically, these representations are describing a conformal minimal immersion of a Riemann surface M into the three-dimensional Euclidean space $\mathbb{R}^3$. Precisely, given a pair (h, g) consisting of a holomorphic and a meromorphic function, the formula

$$f(x, y) = \Re\mathrm{e}\int((1-g^2(z))h(z), (1+g^2(z))h(z), 2g(z)h(z))\,dz,$$

with $z = x + iy$ some complex coordinate, gives a local parametrization of a minimal surface in Euclidean three-space. Conversely every minimal surface can be parametrized in this way with respect to isothermal coordinates. However, relaxing the condition of holomorphicity on the pair (h, g), this representation is much more general and can actually describe all surfaces in $\mathbb{R}^3$ as shown in [Kenmotsu 1979].

This approach was reformulated in a more concise and simpler way in terms of spinor fields by B.G. Konopelchenko [1996], Konopelchenko and I.A. Taĭmanov

MSC2010: 53C27, 53C40.
Keywords: Spin geometry, isometric immersions, Weierstrass representation.

[1996], Taĭmanov [1997a] and R. Kusner and N. Schmitt [1996]. These so-called spinorial Weierstrass representations were studied extensively by these authors and many others, in dimension 3 and 4 (see [Taĭmanov 1997b; Konopelchenko 2000; Konopelchenko and Landolfi 1999; 2000] and the references there).

However these formulae were given in local coordinates and remained purely computational until Friedrich [1998] gave an elegant and geometrically invariant description using spinor bundles. We point out that the equivalence between the two approaches was recently showed in [Romon and Roth 2013]. The main idea is to use the identification between the ambient spinor bundle restricted to the surface and the spinor bundle of the surface. Note that the condition to be a spin manifold is not restrictive here since any oriented surface is also spin. More generally, the restriction φ of a parallel spinor field on $\mathbb{R}^{n+1}$ to an oriented Riemannian hypersurface M^n is a solution of a *generalized Killing equation*

$$\nabla_X^{\Sigma M}\varphi = A(X)\cdot\varphi,$$

where $\nabla^{\Sigma M}$ and $\cdot$ are respectively the spin connection and the Clifford multiplication on M, and A is the shape operator of the immersion. Conversely, Friedrich showed that in the case where M is a simply connected surface, if there exists a particular spinor field φ satisfying the generalized Killing equation, where A is an arbitrary field of symmetric endomorphisms of the tangent bundle, then there exists an isometric immersion of M into $\mathbb{R}^3$ with shape operator A. Moreover, φ is the restriction to M of a parallel spinor of $\mathbb{R}^3$. The proof consists of showing that A indeed satisfies the structure equations. This result was generalized to surfaces into other three-dimensional ambient spaces [Morel 2005; Nakad and Roth 2012; Roth 2010; Taĭmanov 2004], to three-dimensional manifolds into four-dimensional space forms [Lawn and Roth 2010; Nakad and Roth 2012] and also to the two-dimensional pseudo-Riemannian setting [Lawn and Roth 2011]. However the question whether a generalized Killing spinor on a manifold of arbitrary dimension gives rise to an isometric immersion into some Euclidean space remained until now unanswered. Some of the few achievements in this direction were obtained in [Ammann et al. 2013] for real analytic manifolds and in [Bär et al. 2005; Nakad 2011] when A is a Codazzi tensor, showing the existence of an immersion into a Ricci flat manifold admitting a parallel spinor which restricts to φ.

Similarly, in higher codimension, very little is known. In [Bayard et al. 2013], we extended the approach to the case of surfaces in four-dimensional space forms. The key point was to use the remark due to Bär [1998] that an ambient spinor restricted to an immersed submanifold M can be identified with a section of the spinor bundle of the submanifold twisted with the spin bundle of the normal bundle. This was then extended to the pseudo-Riemannian setting in [Bayard 2013; Bayard and Patty 2015].

Following the same idea, we use in this paper a particular twisted spin bundle over a spin manifold of arbitrary dimension to give a geometrically invariant spinorial representation of submanifolds of Euclidean spaces in any codimension. Note that our proof does not use the structure equations but merely the existence of a special generalized Killing spinor on the manifold. Precise definitions are given in the first sections of the paper. We later show that one indeed recovers the previously mentioned result of Friedrich [1998] in the case of surfaces in $\mathbb{R}^3$, as well as the one of Lawn and Roth [2010] for three-dimensional hypersurfaces and of Bayard, Lawn and Roth [Bayard et al. 2013] for surfaces in $\mathbb{R}^4$ (Section 7). It is worth pointing out that the study of generalized Killing spinors has revealed very interesting applications. Moroianu and Semmelmann [2014] were for instance able to construct new examples of Lagrangian submanifolds of the nearly Kähler $\mathbb{S}^3 \times \mathbb{S}^3$ using the existence of such spinors on the sphere $\mathbb{S}^3$. Moreover it is well known that there is a close relationship to G-structures: for instance a generalized Killing spinor defines a cocalibrated G_2-structure on the manifold in dimension 7 and a half-flat SU(3)-structure in dimension 6 (see for example [Chiossi and Salamon 2002]). However the existence of such spinors is a nontrivial problem: our construction is therefore of particular interest.

Besides the above mentioned, we discuss several other applications. A notable achievement is a new and concise proof of the fundamental theorem of submanifold theory. In the special case of surfaces, we show that our approach is equivalent to the spinorial Weierstrass representations, i.e., we obtain explicit formulae in terms of functions involving the components of the spinor field which are holomorphic if the surface is minimal. Our result can thus be seen as a generalization of most of the concrete Weierstrass representation formulae existing in the literature: it provides a general framework to understand formulae appearing in a variety of contexts. Moreover, since the basic ideas and constructions behind our representation are fairly simple, we hope that our result will be useful to obtain new concrete Weierstrass representation formulae, once some geometric context is specified: this is especially interesting for surfaces, in low-dimensional pseudo-Riemannian space forms, under some curvature assumptions.

Finally, in the last section, we extend our result to submanifolds immersed into the other space forms $\mathbb{S}^n$ and $\mathbb{H}^n$, and recover the results of Morel [2005] and Taĭmanov [2004] if $n = 3$.

2. Preliminaries

The spin representation. Let us denote by Cl_n the real Clifford algebra on $\mathbb{R}^n$ with its standard scalar product. We consider the representation

$$\rho : \mathrm{Cl}_n \to \mathrm{End}(\mathrm{Cl}_n), \quad a \mapsto (\xi \mapsto a\xi)$$

and its restriction to the group $\mathrm{Spin}(n)$

$$\rho_{|\,\mathrm{Spin}(n)} : \mathrm{Spin}(n) \to \mathrm{GL}(\mathrm{Cl}_n), \quad a \mapsto (\xi \mapsto a\xi).$$

Note that this is not the adjoint representation of the spin group on the Clifford algebra, but rather the representation given by left multiplication.

Moreover we want to point out that we are not taking as usual the restriction of an irreducible representation of the Clifford algebra to the spin group, but that we consider instead the restriction of the entire real Clifford algebra. This real representation splits into a sum of 2^k copies of spinor spaces of dimension 2^{n-k}, where the number k depends on the dimension n and can be computed using the Radon–Hurwitz numbers (we refer to [Lounesto 2001] for further details).

If $p+q=n$, we have a natural map

$$\mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Spin}(n)$$

associated to the splitting $\mathbb{R}^n = \mathbb{R}^p \oplus \mathbb{R}^q$ and to the corresponding isomorphism of Clifford algebras

$$\mathrm{Cl}_n = \mathrm{Cl}_p \,\hat{\otimes}\, \mathrm{Cl}_q,$$

where $\hat{\otimes}$ denotes the $\mathbb{Z}_2$-graded tensor product. We get thus the following representation, still denoted by ρ,

$$\rho : \mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{GL}(\mathrm{Cl}_n), \quad a \mapsto (\xi \mapsto a\xi). \tag{1}$$

***The twisted spinor bundle* Σ.** We consider M a p-dimensional Riemannian manifold, $E \to M$ a bundle of rank q, with a fiber metric and a compatible connection. We assume that E and TM are oriented and spin, with given spin structures

$$\tilde{Q}_M \xrightarrow{2:1} Q_M \quad \text{and} \quad \tilde{Q}_E \xrightarrow{2:1} Q_E$$

where Q_M and Q_E are the bundles of positively oriented orthonormal frames of TM and E, and we set

$$\tilde{Q} := \tilde{Q}_M \times_M \tilde{Q}_E;$$

this is a $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$-principal bundle. We define the associated bundle

$$\Sigma := \tilde{Q} \times_\rho \mathrm{Cl}_n$$

and its restriction

$$U\Sigma := \tilde{Q} \times_\rho \mathrm{Spin}(n) \subset \Sigma \tag{2}$$

where ρ is the representation (1) given by left multiplication. We remark that if we used the adjoint representation instead, we would just get the Clifford algebra bundle. Again we point out that our spinor bundle Σ is a real vector bundle with fiber the entire Clifford algebra and not, as usual, an irreducible complex Clifford module.

The vector bundle Σ is equipped with the covariant derivative ∇ naturally associated to the spinorial connections on $\tilde{Q}_M$ and $\tilde{Q}_E$.

Remark. The bundle Σ is a spinor bundle on TM twisted by a spinor bundle on E: indeed, let us consider the representations

$$\rho_1 : \mathrm{Spin}(p) \to \mathrm{GL}(\mathrm{Cl}_p) \quad \text{and} \quad \rho_2 : \mathrm{Spin}(q) \to \mathrm{GL}(\mathrm{Cl}_q)$$

given by left multiplication, and the associated bundles

$$\Sigma_1 := \tilde{Q}_M \times_{\rho_1} \mathrm{Cl}_p \quad \text{and} \quad \Sigma_2 := \tilde{Q}_E \times_{\rho_2} \mathrm{Cl}_q$$

equipped with their natural connections ∇^1 and ∇^2; then

$$\Sigma_1 \otimes \Sigma_2 \simeq \Sigma \quad \text{and} \quad \nabla^1 \otimes \mathrm{id}_{\Sigma_2} \oplus \mathrm{id}_{\Sigma_1} \otimes \nabla^2 \simeq \nabla.$$

This is a consequence of the fact that the natural isomorphism

$$i : \mathrm{Cl}_p \otimes \mathrm{Cl}_q \xrightarrow{\sim} \mathrm{Cl}_n, \quad \xi_1 \otimes \xi_2 \mapsto \xi_1 \xi_2$$

is an equivalence of representations of $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$, i.e., for $g_1 \in \mathrm{Spin}(p)$ and $g_2 \in \mathrm{Spin}(q)$,

$$i \circ \rho_1(g_1) \otimes \rho_2(g_2) = \rho(g_1, g_2) \circ i;$$

indeed, if $\xi_1 \in \mathrm{Cl}_p$ and $\xi_2 \in \mathrm{Cl}_q$,

$$\begin{aligned} i(\rho_1(g_1) \otimes \rho_2(g_2)(\xi_1 \otimes \xi_2)) = i(g_1\xi_1 \otimes g_2\xi_2) &= g_1\xi_1 g_2\xi_2 \\ &= g_1 g_2 \xi_1 \xi_2 = \rho(g_1, g_2)(i(\xi_1 \otimes \xi_2)), \end{aligned}$$

where the products in the third and fourth terms are products in Cl_n (note that ξ_1 and g_2 commute since ξ_1 belongs to Cl_p and g_2 is a product of an even number of vectors belonging to $\mathbb{R}^q$).

As in the usual construction in spin geometry, the spin bundle Σ is endowed with a natural action of the Clifford bundle $\mathrm{Cl}(TM \oplus E)$: indeed, the Clifford product

$$\mathrm{Cl}(\mathbb{R}^p \oplus \mathbb{R}^q) \times \mathrm{Cl}_n \to \mathrm{Cl}_n, \quad (\eta, \xi) \mapsto \eta \cdot \xi$$

is $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$ equivariant, if the action of $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$ on $\mathrm{Cl}(\mathbb{R}^p \oplus \mathbb{R}^q)$ is the adjoint action, and the action on Cl_n is the left multiplication: we obviously have, for $(g_1, g_2) \in \mathrm{Spin}(p) \times \mathrm{Spin}(q)$ and $g = g_1 g_2 \in \mathrm{Spin}(n)$,

$$(g\xi g^{-1}) \cdot (g\eta) = g \cdot (\xi\eta) \quad \text{for } \xi, \eta \in \mathrm{Cl}_n\,.$$

A $\mathbf{Cl}_n$***-valued bilinear map on*** $\boldsymbol{\Sigma}$. Let us denote by $\tau : \mathrm{Cl}_n \to \mathrm{Cl}_n$ the antiautomorphism of Cl_n such that

$$\tau(x_1 \cdot x_2 \cdots x_k) = x_k \cdots x_2 \cdot x_1 \quad \text{for all } x_1, x_2, \dots, x_k \in \mathbb{R}^n,$$

where '$\cdot$' denotes as usual the Clifford multiplication, and set

$$\langle\langle \cdot\,, \cdot \rangle\rangle : \mathrm{Cl}_n \times \mathrm{Cl}_n \to \mathrm{Cl}_n, \quad (\xi, \xi') \mapsto \tau(\xi')\xi. \tag{3}$$

This map is $\mathrm{Spin}(n)$-invariant: for all $\xi, \xi' \in \mathrm{Cl}_n$ and $g \in \mathrm{Spin}(n)$ we have

$$\langle\langle g\xi, g\xi' \rangle\rangle = \tau(g\xi')g\xi = \tau(\xi')\tau(g)g\xi = \tau(\xi')\xi = \langle\langle \xi, \xi' \rangle\rangle,$$

since $\mathrm{Spin}(n) \subset \{g \in \mathrm{Cl}_n^0 : \tau(g)g = 1\}$; this map thus induces a Cl_n-valued map

$$\langle\langle \cdot\,, \cdot \rangle\rangle : \Sigma \times \Sigma \to \mathrm{Cl}_n, \quad (\varphi, \varphi') \mapsto \langle\langle [\varphi], [\varphi'] \rangle\rangle \tag{4}$$

where $[\varphi]$ and $[\varphi'] \in \mathrm{Cl}_n$ represent φ and φ' in some spinorial frame $\tilde{s} \in \tilde{Q}$.

Lemma 2.1. *The map* $\langle\langle \cdot\,, \cdot \rangle\rangle$: $\Sigma \times \Sigma \to \mathrm{Cl}_n$ *satisfies the following properties*: *for all* $\varphi, \psi \in \Gamma(\Sigma)$ *and* $X \in \Gamma(TM)$,

$$\langle\langle \varphi, \psi \rangle\rangle = \tau\langle\langle \psi, \varphi \rangle\rangle \tag{5}$$

and

$$\langle\langle X \cdot \varphi, \psi \rangle\rangle = \langle\langle \varphi, X \cdot \psi \rangle\rangle. \tag{6}$$

Proof. We have

$$\langle\langle \varphi, \psi \rangle\rangle = \tau[\psi][\varphi] = \tau(\tau[\varphi][\psi]) = \tau\langle\langle \psi, \varphi \rangle\rangle$$

and

$$\langle\langle X \cdot \varphi, \psi \rangle\rangle = \tau[\psi][X][\varphi] = \tau([X][\psi])[\varphi] = \langle\langle \varphi, X \cdot \psi \rangle\rangle,$$

where $[\varphi]$, $[\psi]$ and $[X] \in \mathrm{Cl}_n$ represent φ, ψ and X in some given frame $\tilde{s} \in \tilde{Q}$. $\square$

Lemma 2.2. *The connection* ∇ *is compatible with the product* $\langle\langle \cdot\,, \cdot \rangle\rangle$:

$$\partial_X\langle\langle \varphi, \varphi' \rangle\rangle = \langle\langle \nabla_X\varphi, \varphi' \rangle\rangle + \langle\langle \varphi, \nabla_X\varphi' \rangle\rangle$$

for all $\varphi, \varphi' \in \Gamma(\Sigma)$ *and* $X \in \Gamma(TM)$.

Proof. If $\varphi = [\tilde{s}, [\varphi]]$ is a section of $\Sigma = \tilde{Q} \times_\rho \mathrm{Cl}_n$, we have

$$\nabla_X\varphi = [\tilde{s}, \partial_X[\varphi] + \rho_*(\tilde{s}^*\alpha(X))([\varphi])] \quad \text{for all } X \in TM,$$

where ρ is the representation (1) and α is the connection form on $\tilde{Q}$; the term $\rho_*(\tilde{s}^*\alpha(X))$ is an endomorphism of Cl_n given by the multiplication on the left by an element belonging to $\Lambda^2\mathbb{R}^n \subset \mathrm{Cl}_n$, still denoted by $\rho_*(\tilde{s}^*\alpha(X))$. Such an element satisfies

$$\tau(\rho_*(\tilde{s}^*\alpha(X))) = -\rho_*(\tilde{s}^*\alpha(X)),$$

and we have

$$\begin{aligned}\langle\langle\nabla_X\varphi,\varphi'\rangle\rangle+\langle\langle\varphi,\nabla_X\varphi'\rangle\rangle &= \tau\{[\varphi']\}(\partial_X[\varphi]+\rho_*(\tilde{s}^*\alpha(X))[\varphi]) \\ &\qquad +\tau\{\partial_X[\varphi']+\rho_*(\tilde{s}^*\alpha(X))[\varphi']\}[\varphi] \\ &= \tau\{[\varphi']\}\partial_X[\varphi] \ +\tau\{\partial_X[\varphi']\}[\varphi] \\ &= \partial_X\langle\langle\varphi,\varphi'\rangle\rangle. \qquad \square\end{aligned}$$

3. The spin geometry of a submanifold in $\mathbb{R}^n$

We keep the notation of the previous section, assuming moreover here that M is a submanifold of $\mathbb{R}^n$ and that $E\to M$ is its normal bundle. Let as before $\tilde{Q}_M \xrightarrow{2:1} Q_M$ be a spin structure of M. Our goal is to construct $\tilde{Q}$ such that we obtain an identification

$$\Sigma=\tilde{Q}\times_\rho \mathrm{Cl}_n \simeq M\times \mathrm{Cl}_n\,.$$

Although this type of result is used in several places in the literature, we could not find a complete statement or proof. Therefore we will give a detailed proof, which we believe may be useful in its own right.

Let $(e_1,\dots,e_p)$ resp. $(e_{p+1},\dots,e_{p+q})$ be orthonormal frames of TM resp. E and $Q_{\mathbb{R}^n}$ the bundle of positively oriented orthonormal frames of $\mathbb{R}^n$. We can consider the map

$$\begin{aligned}\iota: Q_M\times_M Q_E &\to Q_{\mathbb{R}^n}\\ ((e_1,\dots,e_p),(e_{p+1},\dots,e_{p+q})) &\mapsto (e_1,e_2,\dots,e_{p+q})\end{aligned}$$

given by the concatenation of frames.

The map

$$\tilde{Q}_M\times_M Q_E\to Q_M\times_M Q_E$$

is obviously a two-to-one covering of $Q_M\times_M Q_E$.

Let now $\tilde{Q}_{\mathbb{R}^n}\xrightarrow{2:1} Q_{\mathbb{R}^n}$ be the (unique) spin structure of $\mathbb{R}^n$. Then the bundle

$$\tilde{Q}:=(\tilde{Q}_M\times_M Q_E)\times_{Q_{\mathbb{R}^n}}\tilde{Q}_{\mathbb{R}^n}$$

is a $\mathrm{Spin}(p)\times\mathrm{Spin}(q)$-principal bundle over M and a four-to-one covering of $Q_M\times_M Q_E$. Observe that $\tilde{Q}=\tilde{Q}_M\times_M\tilde{Q}_E$, where $\tilde{Q}_E:=\tilde{Q}/\mathrm{Spin}(p)$ (and the projection $\tilde{Q}/\mathrm{Spin}(q)\to\tilde{Q}_M$ is a map of principal $\mathrm{Spin}(p)$-bundles, hence an isomorphism). Moreover, $\tilde{Q}_E$ is a spin structure on E, canonically associated to the spin structures on M and $\mathbb{R}^n$.

Claim. *Consider the bundle*

$$\tilde{Q}\times_c\mathrm{Spin}(p+q):=(\tilde{Q}\times\mathrm{Spin}(p+q))/(\mathrm{Spin}(p)\times\mathrm{Spin}(q)),$$

where

$$c : \mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Spin}(p+q)$$

is the map corresponding to the isomorphism of Clifford algebras $\mathrm{Cl}_p \hat{\otimes} \mathrm{Cl}_q \cong \mathrm{Cl}_{p+q}$. *Then there is a canonical isomorphism of* $\mathrm{Spin}(n)$*-principal bundles*,

$$\tilde{Q} \times_c \mathrm{Spin}(p+q) \cong \tilde{Q}_{\mathbb{R}^n}|_M.$$

Proof. Consider the projection $\pi : \tilde{Q} \to \tilde{Q}_{\mathbb{R}^n}$ to the last factor. Then the map

$$\tilde{\pi} : \tilde{Q} \times \mathrm{Spin}(p+q) \to \tilde{Q}_{\mathbb{R}^n}, \quad (\tilde{q}, s) \mapsto s\pi(\tilde{q})$$

satisfies $\tilde{\pi}(s_0\tilde{q}, ss_0^{-1}) = \tilde{\pi}(\tilde{q}, s)$ for any $s_0 \in \mathrm{Spin}(p) \times \mathrm{Spin}(q)$, so $\tilde{\pi}$ descends to a map $\tilde{Q} \times_c \mathrm{Spin}(p+q) \to \tilde{Q}_{\mathbb{R}^n}$. The source is clearly a $\mathrm{Spin}(p+q)$-principal bundle on M, as is the target, and the map is $\mathrm{Spin}(p+q)$-equivariant and the identity over M. Hence it is an isomorphism of principal bundles. □

Corollary 1. *If now* $\rho : \mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{GL}(\mathrm{Cl}_n)$ *is the map given by* $\tilde{\rho} \circ c$, *where* $\tilde{\rho} : \mathrm{Spin}(n) \to \mathrm{GL}(\mathrm{Cl}_n)$ *is the representation induced by left multiplication*, *we get*

$$\tilde{Q} \times_\rho \mathrm{Cl}_n \cong \tilde{Q}_{\mathbb{R}^n}|_M \times_{\tilde{\rho}} \mathrm{Cl}_n \cong M \times \mathrm{Cl}_n .$$

Proof. The first isomorphism is immediate from the claim, and the second follows since $\tilde{Q}_{\mathbb{R}^n}$ is trivial. □

Two connections are thus defined on Σ, the connection ∇ introduced in the previous section and the trivial connection ∂; they satisfy the following Gauss formula:

$$\partial_X \varphi = \nabla_X \varphi + \frac{1}{2} \sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot \varphi \tag{7}$$

for all $\varphi \in \Gamma(\Sigma)$ and all $X \in \Gamma(TM)$, where $B : TM \times TM \to E$ is the second fundamental form of M into $\mathbb{R}^n$. We refer to [Bär 1998] for the proof (in a slightly different context).

4. Spinorial representation of submanifolds in $\mathbb{R}^n$

We state the main result of the paper. Let M be a p-dimensional Riemannian manifold and $E \to M$ a bundle of rank q, with a fiber metric and a compatible connection; we assume that E and TM are oriented and spin, with given spin structures. We keep the notation of Section 2.

Theorem 2. *We moreover assume that M is simply connected*, *and suppose that* $B : TM \times TM \to E$ *is bilinear and symmetric. The following statements are equivalent*:

(1) *There exists a section* $\varphi \in \Gamma(U\Sigma)$ *such that*

$$\nabla_X \varphi = -\frac{1}{2}\sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot \varphi \quad \textit{for all } X \in TM. \tag{8}$$

(2) *There exists an isometric immersion* $F : M \to \mathbb{R}^n$ *with normal bundle* E *and second fundamental form* B. *Moreover*, $F = \int \xi$ *where* ξ *is the* $\mathbb{R}^n$*-valued* 1*-form defined by*

$$\xi(X) := \langle\langle X \cdot \varphi, \varphi \rangle\rangle \quad \textit{for all } X \in TM. \tag{9}$$

The representation formula (9) generalizes the classical Weierstrass representation formula and most of the spinor representation formulae in the literature. Special cases will be studied in Sections 7 and 8.

Remark. Formula (9) also presents the advantage of unifying previously known formulae, and therefore explaining apparent discrepancies in these results. In particular, it is known that the number of spinor fields needed to represent immersions, and also the normalization of these spinor fields, vary depending on the geometric context; for instance, a single spinor field is needed to represent surfaces in $\mathbb{R}^3$ but two spinor fields are necessary to represent general hypersurfaces in dimension 4 [Lawn and Roth 2010], and the required normalization to represent surfaces in $\mathbb{R}^{1,2}$, $\mathbb{R}^{1,3}$ or $\mathbb{R}^{2,2}$ differ to that in $\mathbb{R}^3$ or $\mathbb{R}^4$ [Bayard 2013; Bayard et al. 2013; Bayard and Patty 2015; Friedrich 1998; Lawn 2008]. This is now easily explained by the fact that the usual spinor representation has in general to be replaced by a representation on the Clifford algebra (which is a sum of usual real spinorial representations), and that the convenient normalization is in fact determined by the bundle $U\Sigma$ (whose fiber is the spin group).

Remark. Taking the trace of (8) we get

$$D\varphi = \tfrac{1}{2} p \vec{H} \cdot \varphi,$$

where $D\varphi = \sum_{j=1}^{p} e_j \cdot \nabla_{e_j} \varphi$, and $\vec{H} = (1/p) \sum_{j=1}^{p} B(e_j, e_j)$ is the mean curvature vector of M in $\mathbb{R}^n$. This Dirac equation is known to be equivalent to (8) only for $p = 2$ or 3 (see, e.g., [Friedrich 1998; Lawn and Roth 2010; 2011; Bayard et al. 2013).

Proof. (2) $\Rightarrow$ (1) is a direct consequence of the Gauss formula (7) for a submanifold of $\mathbb{R}^n$: the restriction of parallel spinor fields of the ambient space $\mathbb{R}^n$ to the submanifold M are obviously solutions of equation (8) (recall that in the paper the spinors are constructed with the whole Clifford algebra). The immersion takes the form $F = \int \xi$ where ξ is given by (9) for the special choice $\varphi = 1_{\mathrm{Cl}_n|M}$, since, in that case, for all $X \in TM$,

$$\xi(X) = \tau[\varphi][X][\varphi] = [X] \simeq X,$$

where $[\varphi] = \pm 1_{\mathrm{Cl}_n}$ and $[X] \in \mathbb{R}^n$ represent φ and X in one of the two spinorial frames of $\mathbb{R}^n$ above the canonical basis.

(1) $\Rightarrow$ (2): We will prove that the 1-form ξ defined in (9) indeed gives us an immersion preserving the metric, the second fundamental form and the normal connection. This follows directly from Propositions 4.1 and 4.2 below. □

Proposition 4.1. *Assume that $\varphi \in \Gamma(U\Sigma)$ is a solution of* (8) *and define ξ by* (9). *Then*

(1) *ξ takes its values in $\mathbb{R}^n \subset \mathrm{Cl}_n$;*

(2) *ξ is a closed 1-form: $d\xi = 0$.*

Proof. (1) By the very definition of ξ, we have

$$\xi(X) = \tau[\varphi][X][\varphi] \quad \text{for all } X \in TM,$$

where $[X]$ and $[\varphi]$ represent X and φ in a given frame $\tilde{s}$ of $\tilde{Q}$. Since $[X]$ belongs to $\mathbb{R}^n \subset \mathrm{Cl}_n$ and $[\varphi]$ is an element of $\mathrm{Spin}(n)$, $\xi(X)$ belongs to $\mathbb{R}^n$.

(2) We compute, for $X, Y \in \Gamma(TM)$ such that $\nabla X = \nabla Y = 0$ at some point x_0,

$$\begin{aligned}
\partial_X \xi(Y) &= \langle\langle Y \cdot \nabla_X \varphi, \varphi \rangle\rangle + \langle\langle Y \cdot \varphi, \nabla_X \varphi \rangle\rangle \\
&= (\mathrm{id} + \tau)\langle\langle Y \cdot \varphi, \nabla_X \varphi \rangle\rangle \\
&= (\mathrm{id} + \tau)\left\langle\!\!\left\langle \varphi, -\frac{1}{2}\sum_{j=1}^{p} Y \cdot e_j \cdot B(X, e_j) \cdot \varphi \right\rangle\!\!\right\rangle.
\end{aligned}$$

Hence

$$d\xi(X, Y) = \partial_X \xi(Y) - \partial_Y \xi(X) = (\mathrm{id} + \tau)\langle\langle \varphi, \mathcal{C} \cdot \varphi \rangle\rangle$$

with

$$\mathcal{C} := -\frac{1}{2}\sum_{j=1}^{p} \{Y \cdot e_j \cdot B(X, e_j) - X \cdot e_j \cdot B(Y, e_j)\}.$$

Now, for $X = \sum_{1 \le k \le p} x_k e_k$ and $Y = \sum_{1 \le k \le p} y_k e_k$,

$$\sum_{j=1}^{p} X \cdot e_j \cdot B(Y, e_j) = -B(Y, X) + \sum_{j=1}^{p} \sum_{k \neq j} x_k e_k \cdot e_j \cdot B(Y, e_j)$$

and

$$\sum_{j=1}^{p} Y \cdot e_j \cdot B(X, e_j) = -B(X, Y) + \sum_{j=1}^{p} \sum_{k \neq j} y_k e_k \cdot e_j \cdot B(X, e_j),$$

which yields the formula

$$\mathcal{C} = -\frac{1}{2}\sum_{j=1}^{p}\sum_{k\neq j} e_k \cdot e_j \cdot (y_k B(X, e_j) - x_k B(Y, e_j)).$$

This shows that $\tau[\mathcal{C}] = -[\mathcal{C}]$, which implies that

$$\tau\langle\langle\varphi, \mathcal{C}\cdot\varphi\rangle\rangle = \tau(\tau[\varphi]\tau[\mathcal{C}][\varphi]) = -\tau[\varphi]\tau[\mathcal{C}][\varphi] = -\langle\langle\varphi, \mathcal{C}\cdot\varphi\rangle\rangle.$$

Thus

$$d\xi(X, Y) = (\mathrm{id} + \tau)\langle\langle\varphi, \mathcal{C}\cdot\varphi\rangle\rangle = 0. \qquad \square$$

We keep the notation of Proposition 4.1, and moreover assume that M is simply connected; since ξ is closed by Proposition 4.1 we can consider

$$F : M \to \mathbb{R}^n$$

such that $dF = \xi$. The next proposition follows from the properties of the Clifford product:

Proposition 4.2. (1) *The map $F : M \to \mathbb{R}^n$ is an isometry.*

(2) *The map*

$$\Phi_E : E \to M \times \mathbb{R}^n, \quad X \in E_m \mapsto (F(m), \xi(X))$$

is an isometry between E and the normal bundle of $F(M)$ into $\mathbb{R}^n$, preserving connections and second fundamental forms. Here, for $X \in E$, $\xi(X)$ still stands for the quantity $\langle\langle X\cdot\varphi, \varphi\rangle\rangle$.

Proof. For $X, Y \in \Gamma(TM \oplus E)$, we have

$$\begin{aligned}\langle\xi(X), \xi(Y)\rangle &= -\tfrac{1}{2}(\xi(X)\xi(Y) + \xi(Y)\xi(X))\\ &= -\tfrac{1}{2}(\tau[\varphi][X][\varphi]\tau[\varphi][Y][\varphi] + \tau[\varphi][Y][\varphi]\tau[\varphi][X][\varphi])\\ &= -\tfrac{1}{2}\tau[\varphi]([X][Y] + [Y][X])[\varphi]\\ &= \langle X, Y\rangle,\end{aligned}$$

since $[X][Y] + [Y][X] = -2\langle[X], [Y]\rangle = -2\langle X, Y\rangle$. This implies that F is an isometry, and that Φ_E is a bundle map between E and the normal bundle of $F(M)$ into $\mathbb{R}^n$ which preserves the metrics of the fibers. Let us denote by B_F and ∇'^F the second fundamental form and the normal connection of the immersion F; we want to show that

$$\xi(B(X, Y)) = B_F(\xi(X), \xi(Y)) \quad \text{and} \quad \xi(\nabla'_X N) = \nabla'^F_{\xi(X)}\xi(N) \tag{10}$$

for $X, Y \in \Gamma(TM)$ and $N \in \Gamma(E)$. First,

$$B_F(\xi(X), \xi(Y)) = \{\partial_X \xi(Y)\}^N,$$

where the superscript N means that we consider the component of the vector which is normal to the immersion. We showed in the proof of Proposition 4.1 that fixing a point $x_0 \in M$, and assuming that $\nabla Y = 0$ at x_0 we have

$$\partial_X \xi(Y) = -\frac{1}{2}(\mathrm{id}+\tau)\left\langle\!\!\left\langle \varphi, Y \cdot \sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot \varphi \right\rangle\!\!\right\rangle,$$

and that moreover

$$Y \cdot \sum_{j=1}^{p} e_j \cdot B(X, e_j) = -B(X, Y) + \mathcal{D},$$

where $\mathcal{D}$ is a term which satisfies $\tau\mathcal{D} = -\mathcal{D}$. This implies that

$$B_F(\xi(X), \xi(Y)) = \left\{\tfrac{1}{2}(\mathrm{id}+\tau)\langle\langle \varphi, B(X, Y) \cdot \varphi \rangle\rangle\right\}^N = \xi(B(X, Y)),$$

where the last equality holds since $\tau[B(X, Y)] = [B(X, Y)]$ and $\xi(B(X, Y))$ is normal to the immersion. We finally show the second identity in (10): we have

$$\nabla'^F_{\xi(X)} \xi(N) = (\partial_X \xi(N))^N = \langle\langle \nabla'_X N \cdot \varphi, \varphi \rangle\rangle^N + \langle\langle N \cdot \nabla_X \varphi, \varphi \rangle\rangle^N + \langle\langle N \cdot \varphi, \nabla_X \varphi \rangle\rangle^N.$$

The first term in the right-hand side is $\xi(\nabla'_X N)$, and we only need to show that

$$\langle\langle N \cdot \nabla_X \varphi, \varphi \rangle\rangle^N + \langle\langle N \cdot \varphi, \nabla_X \varphi \rangle\rangle^N = 0. \tag{11}$$

We have

$$\begin{aligned}
\langle\langle N \cdot \nabla_X \varphi, \varphi \rangle\rangle + \langle\langle N \cdot \varphi, \nabla_X \varphi \rangle\rangle &= (\mathrm{id}+\tau)\langle\langle N \cdot \nabla_X \varphi, \varphi \rangle\rangle \\
&= \frac{1}{2}(\mathrm{id}+\tau)\left\langle\!\!\left\langle \sum_{j=1}^{p} e_j \cdot N \cdot B(X, e_j) \cdot \varphi, \varphi \right\rangle\!\!\right\rangle,
\end{aligned}$$

and the identity (11) will thus be proved if we show that this vector is tangent to the immersion. We have

$$\begin{aligned}
\sum_{j=1}^{p} e_j \cdot N \cdot B(X, e_j) &= -\sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot N - 2\sum_{j=1}^{p} \langle B(X, e_j), N \rangle e_j \\
&= -\sum_{j=1}^{p} B(X, e_j) \cdot N \cdot e_j - 2B^*(X, N) \\
&= -\tau\left(\sum_{j=1}^{p} e_j \cdot N \cdot B(X, e_j)\right) - 2B^*(X, N),
\end{aligned}$$

where we have set $B^*(X,N)=\sum_{j=1}^{p}\langle B(X,e_j),N\rangle e_j$; thus

$$\frac{1}{2}(\mathrm{id}+\tau)\left\langle\!\!\left\langle\sum_{j=1}^{p} e_j\cdot N\cdot B(X,e_j)\cdot\varphi,\varphi\right\rangle\!\!\right\rangle=-\langle\langle B^*(X,N)\cdot\varphi,\varphi\rangle\rangle,$$

which is a vector tangent to the immersion since $B^*(X,N)$ belongs to TM; (11) follows, which finishes the proof. □

Remark. The group $\mathrm{Spin}(n)$ naturally acts on $U\Sigma$ by multiplication on the right, and if $\varphi\in\Gamma(U\Sigma)$ is a solution of (8) and g_0 belongs to $\mathrm{Spin}(n)$, then $\varphi\cdot g_0$ is also a solution of (8); in fact, $\varphi\cdot g_0$ defines an immersion which is congruent to the immersion defined by φ: indeed, for all $X\in\Gamma(TM)$,

$$\xi_{\varphi\cdot g_0}(X)=\tau[\varphi\cdot g_0][X][\varphi\cdot g_0]=\tau(g_0)\tau[\varphi][X][\varphi]g_0=\tau(g_0)\xi_\varphi(X)g_0,$$

i.e.,

$$\xi_{\varphi.g_0}=\mathrm{Ad}({g_0}^{-1})\circ\xi_\varphi;$$

the linear part of the rigid motion between the immersions defined by φ and $\varphi\cdot g_0$ is thus $\mathrm{Ad}({g_0}^{-1})\in\mathrm{SO}(n)$.

5. An application: the fundamental theorem of submanifold theory

We first recall the equations of Gauss, Ricci and Codazzi for the symmetric bilinear form B. Let R^T and R^N stand respectively for the curvature tensors of the connections on TM and on E. Further, let $B^*:TM\times E\to TM$ be the bilinear map such that for all $X,Y\in\Gamma(TM)$ and $N\in\Gamma(E)$

$$\langle B(X,Y),N\rangle=\langle Y,B^*(X,N)\rangle.$$

Then we have, for all $X,Y,Z\in\Gamma(TM)$ and $N\in\Gamma(E)$,

(1) the Gauss equation

$$R^T(X,Y)Z=B^*(X,B(Y,Z))-B^*(Y,B(X,Z)),$$

(2) the Ricci equation

$$R^N(X,Y)N=B(X,B^*(Y,N))-B(Y,B^*(X,N)),$$

(3) the Codazzi equation

$$\tilde\nabla_X B(Y,Z)=\tilde\nabla_Y B(X,Z);$$

in the last equation, $\tilde\nabla$ denotes the natural connection on $T^*M\otimes T^*M\otimes E$.

Proposition 5.1. *The equations of Gauss, Ricci and Codazzi on B are the integrability conditions of* (8).

Proof. We assume that $\varphi \in \Gamma(U\Sigma)$ is a solution of (8) and compute the curvature

$$R(X,Y)\varphi = \nabla_X\nabla_Y\varphi - \nabla_Y\nabla_X\varphi - \nabla_{[X,Y]}\varphi.$$

We fix a point $x_0 \in M$, and assume that $\nabla X = \nabla Y = 0$ at x_0. We have

$$\begin{aligned}\nabla_X\nabla_Y\varphi &= -\frac{1}{2}\sum_{j=1}^{p} e_j \cdot (\tilde{\nabla}_X B(Y,e_j)\cdot\varphi + B(Y,e_j)\cdot\nabla_X\varphi)\\ &= -\frac{1}{2}\sum_{j=1}^{p} e_j \cdot \tilde{\nabla}_X B(Y,e_j)\cdot\varphi - \frac{1}{4}\sum_{j,k=1}^{p} e_j\cdot e_k\cdot B(Y,e_j)\cdot B(X,e_k).\end{aligned}$$

Thus

$$\begin{aligned}R(X,Y)\varphi = &-\frac{1}{2}\sum_{j=1}^{p} e_j\cdot(\tilde{\nabla}_X B(Y,e_j) - \tilde{\nabla}_Y B(X,e_j))\cdot\varphi\\ &+\frac{1}{4}\sum_{j\neq k} e_j\cdot e_k\cdot(B(X,e_j)\cdot B(Y,e_k) - B(Y,e_j)\cdot B(X,e_k))\cdot\varphi\\ &-\frac{1}{4}\sum_{j=1}^{p}(B(X,e_j)\cdot B(Y,e_j) - B(Y,e_j)\cdot B(X,e_j))\cdot\varphi.\end{aligned}\tag{12}$$

We compute the last two terms in the following lemma:

Lemma 5.2. *Let us set*

$$\mathcal{A} := \frac{1}{4}\sum_{j\neq k} e_j\cdot e_k\cdot(B(X,e_j)\cdot B(Y,e_k) - B(Y,e_j)\cdot B(X,e_k))$$

and

$$\mathcal{B} := -\frac{1}{4}\sum_{j=1}^{p}(B(X,e_j)\cdot B(Y,e_j) - B(Y,e_j)\cdot B(X,e_j)).$$

We have

$$\mathcal{A} = \frac{1}{2}\sum_{j<k}\{\langle B^*(X,B(Y,e_j)),e_k\rangle - \langle B^*(Y,B(X,e_j)),e_k\rangle\}e_j\cdot e_k$$

and

$$\mathcal{B} = \frac{1}{2}\sum_{k<l}\langle B(X,B^*(Y,n_k)) - B(Y,B^*(X,n_k)),n_l\rangle n_k\cdot n_l.$$

Here, $e_1,\ldots,e_p$ and $n_1,\ldots,n_q$ are orthonormal basis of $T_{x_o}M$ and E_{x_o}, respectively.

Proof. For the computation of $\mathcal{A}$, we notice that

$$\sum_{j\neq k} e_j\cdot e_k\cdot B(Y,e_j)\cdot B(X,e_k)=-\sum_{j\neq k} e_j\cdot e_k\cdot B(Y,e_k)\cdot B(X,e_j),$$

and get

$$\begin{aligned}
\mathcal{A}&=\frac{1}{4}\sum_{j\neq k} e_j\cdot e_k\cdot (B(X,e_j)\cdot B(Y,e_k)+B(Y,e_k)\cdot B(X,e_j))\\
&=-\frac{1}{2}\sum_{j\neq k}\langle B(X,e_j),\,B(Y,e_k)\rangle e_j\cdot e_k\\
&=-\frac{1}{2}\sum_{j<k}\{\langle B(X,e_j),\,B(Y,e_k)\rangle-\langle B(Y,e_j),\,B(X,e_k)\rangle\}e_j\cdot e_k\\
&=-\frac{1}{2}\sum_{j<k}\{\langle B^*(Y,\,B(X,e_j)),e_k\rangle-\langle B^*(X,\,B(Y,e_j)),e_k\rangle\}e_j\cdot e_k.
\end{aligned}$$

For the computation of $\mathcal{B}$, we write

$$B(Y,e_j)=\sum_k\langle B(Y,e_j),n_k\rangle n_k\quad\text{and}\quad B(X,e_j)=\sum_l\langle B(X,e_j),n_l\rangle n_l$$

and get

$$\begin{aligned}
\sum_j B(Y,e_j)\cdot B(X,e_j)&=\sum_{kl}\sum_j\langle B(Y,e_j),n_k\rangle\langle B(X,e_j),n_l\rangle n_k\cdot n_l\\
&=\sum_{kl}\sum_j\langle e_j,\,B^*(Y,n_k)\rangle\langle e_j,\,B^*(X,n_l)\rangle n_k\cdot n_l\\
&=\sum_{kl}\langle B^*(Y,n_k),\,B^*(X,n_l)\rangle n_k\cdot n_l\\
&=\sum_{kl}\langle B(X,\,B^*(Y,n_k)),n_l\rangle n_k\cdot n_l;
\end{aligned}$$

thus

$$\begin{aligned}
\mathcal{B}&=\frac{1}{4}\sum_{kl}\langle B(X,\,B^*(Y,n_k))-B(Y,\,B^*(X,n_k)),n_l\rangle n_k\cdot n_l\\
&=\frac{1}{2}\sum_{k<l}\langle B(X,\,B^*(Y,n_k))-B(Y,\,B^*(X,n_k)),n_l\rangle n_k\cdot n_l.
\end{aligned}$$

$\square$

On the other hand, the curvature of the spinorial connection is given by

$$R(X,Y)\varphi=\frac{1}{2}\Bigg(\sum_{1\le j<k\le p}\langle R^T(X,Y)(e_j),e_k\rangle e_j\cdot e_k+\sum_{1\le k<l\le q}\langle R^N(X,Y)(n_k),n_l\rangle n_k\cdot n_l\Bigg)\cdot\varphi. \tag{13}$$

We now compare the expressions (12) and (13) using the calculations in Lemma 5.2: since in a given frame $\tilde{s}$ belonging to $\tilde{Q}$, φ is represented by an element which is invertible in Cl_n (it is in fact represented by an element belonging to $\mathrm{Spin}(n)$), we may identify the coefficients and get

$$\langle R^T(X,Y)(ej),e_k\rangle=\langle B^*(X,B(Y,e_j)),e_k\rangle-\langle B^*(Y,B(X,e_j)),e_k\rangle,$$

$$\langle R^N(X,Y)(n_k),n_l\rangle=\langle B(X,B^*(Y,n_k)),n_l\rangle-\langle B(Y,B^*(X,n_k)),n_l\rangle$$

and

$$\tilde{\nabla}_X B(Y,e_j)-\tilde{\nabla}_Y B(X,e_j)=0$$

for all the indices. These equations are the equations of Gauss, Ricci and Codazzi.

We finally show that the equations of Gauss, Codazzi and Ricci are also sufficient to get a solution of (8): by the computation above, the connection on Σ defined by

$$\nabla'_X\varphi:=\nabla_X\varphi+\frac{1}{2}\sum_{j=1}^{p}e_j\cdot B(X,e_j)\cdot\varphi \tag{14}$$

for all $\varphi\in\Gamma(\Sigma)$ and $X\in\Gamma(TM)$ is then a flat connection. Moreover, this connection may be regarded as a connection on the principal bundle $U\Sigma$ (with the group $\mathrm{Spin}(n)$ acting from the right): indeed, ∇ defines such a connection (since it comes from a connection on $\tilde{Q}$ and by (2)), and the right-hand side term in (14) defines a linear map

$$TM\to\chi_V^{\mathrm{inv}}(U\Sigma),\qquad X\mapsto\varphi\mapsto\frac{1}{2}\sum_{j=1}^{p}e_j\cdot B(X,e_j)\cdot\varphi$$

from TM to the vector fields $\chi_V^{\mathrm{inv}}(U\Sigma)$ on $U\Sigma$ which are vertical and invariant under the action of the group (these vector fields are indeed of the form $\varphi\mapsto\eta\cdot\varphi$, $\eta\in\Lambda^2(TM\oplus E)\subset\mathrm{Cl}(TM\oplus E)$). Since a flat connection on a principal bundle admits a local parallel section, there exists a local section $\varphi\in\Gamma(U\Sigma)$ such that $\nabla'\varphi=0$, and thus a solution of (8). □

As a consequence of Theorem 2 and Proposition 5.1 we therefore get immediately

Corollary 3 (fundamental theorem of submanifold theory). *We keep the hypotheses and notation of Section 2, and moreover assume that M is simply connected and that $B:TM\times TM\to E$ is bilinear, symmetric and satisfies the equations of Gauss,* *Codazzi and Ricci. Then there exists an isometric immersion of M into $\mathbb{R}^n$ with normal bundle E and second fundamental form B. The immersion is unique up to a rigid motion in $\mathbb{R}^n$.*

Proof. As proved in Proposition 5.1, the equations of Gauss, Codazzi and Ricci are exactly the integrability conditions of (8). By Theorem 2, with a solution $\varphi \in \Gamma(U\Sigma)$ of equation (8) at hand, $F = \int \xi$, where ξ is the 1-form defined in (9), is the immersion. Finally, a solution of (8) is unique up to the multiplication on the right by an element of $\mathrm{Spin}(n)$ (since this is a parallel section of the $\mathrm{Spin}(n)$ principal bundle $U\Sigma$, see the proof of Proposition 5.1); the multiplication on the right of φ by an element of $\mathrm{Spin}(n)$ and the adding of a constant vector in $\mathbb{R}^n$ in the last integration give an immersion which is congruent to the immersion defined by φ (see the remark on page 63). □

6. Relation to the Gauss map

We show here that the spinor field representing the immersion is an horizontal lift of the Gauss map. Let us consider the Grassmannian $\mathrm{Gr}_{p,n} \subset \Lambda^p(\mathbb{R}^n)$ of the oriented p-dimensional linear spaces in $\mathbb{R}^n$. Using the natural isomorphism of vector spaces between the exterior algebra over $\mathbb{R}^n$ and Cl_n, $\mathrm{Gr}_{p,n}$ identifies with the set

$$\mathcal{Q}_o = \{e_1 \cdot e_2 \cdots e_p \in \mathrm{Cl}_n, e_i \in \mathbb{R}^n, |e_i| = 1, e_i \perp e_j, i, j = 1, \ldots, p, i \neq j\}.$$

We recall that for an oriented p-dimensional submanifold $F : M \to \mathbb{R}^n$ the Gauss map is defined as the map which assigns each point $x \in M$ to the oriented tangent space $dF(T_xM)$ considered as a vector subspace of $\mathbb{R}^n$. It can hence be seen as the map into the Grassmannian

$$G : M \to \mathcal{Q}_o, \quad x \mapsto dF(e_1) \cdot dF(e_2) \cdots dF(e_p),$$

where $e_1, e_2, \ldots, e_p$ is a positively oriented orthonormal basis of T_xM.

We assume that the immersion F of M into $\mathbb{R}^n$ is given by a spinor field φ, as in Theorem 2.

Proposition 6.1. *The spinor field φ, which is a section of $U\Sigma$, is a lift of the Gauss map: the diagram*

$$\begin{tikzcd} & U\Sigma \arrow[d, "\chi"] \\ M \arrow[ur, bend left, "\varphi"] \arrow[r, "\bar{G}"'] & M \times \mathcal{Q}_o \end{tikzcd}$$

commutes, where $\bar{G}(x) = (x, G(x))$ and the projection $U\Sigma \to M \times \mathcal{Q}_o$ is given by

(15) $$\chi : \varphi \in U\Sigma_x \mapsto (x, \langle\langle \omega \cdot \varphi, \varphi \rangle\rangle),$$

where ω is the volume form in $\mathrm{Cl}(T_xM)$ (*the product of the elements of a positively oriented orthonormal basis of* T_xM).

It is moreover parallel with respect to the connection

$$\nabla'_X\varphi := \nabla_X\varphi + \frac{1}{2}\sum_{j=1}^{p} e_j \cdot B(X, e_j)\cdot\varphi$$

on $U\Sigma$.

Proof. We first explain why the map χ as defined indeed has target $M \times \mathcal{Q}_o$. Consider the map

$$\Xi : \Sigma \times_M \mathrm{Cl}(TM) \to M \times \mathrm{Cl}_n, \quad (\psi, c) \mapsto \langle\langle c\cdot\psi, \psi\rangle\rangle =: \Xi_\psi(c).$$

Suppose $\psi \in U\Sigma$ and $c = e_1 \cdots e_k$ for k orthonormal vectors $e_1, \dots, e_k \in T_xM$. Then, we can rewrite $\Xi_\psi(c) = \langle\langle c\cdot\psi, \psi\rangle\rangle$ in any spinorial frame at x as

$$\tau[\psi][e_1]\cdots[e_k][\psi] = (\tau[\psi][e_1][\psi])(\tau[\psi][e_2][\psi])\cdots(\tau[\psi][e_k][\psi]). \tag{16}$$

The k vectors on the right-hand side are still orthonormal, so $\Xi_\psi(c)$ lies in the corresponding Grassmannian $\mathrm{Gr}_{k,n}$. Consequently $\chi(\psi) = \Xi_\psi(\omega)$ lies in $M \times \mathcal{Q}_o$.

We next verify the formula for the Gauss map. Recall that the immersion is given by $F = \int \xi$, where ξ is the 1-form defined by $\xi(X) = \langle\langle X\cdot\varphi, \varphi\rangle\rangle$ for all $X \in TM$. Thus, $dF = \xi$. We fix a positively oriented and orthonormal frame $(e_1, \dots, e_p)$ of TM, and a spinorial frame $\tilde{s} \in \tilde{Q}$ which is above $(e_1, \dots, e_p)$. Then, $\omega = e_1\cdots e_p$. In any spinorial frame, $\tau[\varphi][v][\varphi] = \xi(v)$ for all $v \in T_xM$. Therefore (16) yields that $\chi(\varphi) = \xi(e_1)\xi(e_2)\cdots\xi(e_p) = G(x)$. This proves the first part of the proposition.

Finally, φ is horizontal with respect to ∇' since it is a solution of (8). □

7. Special cases: minimal surfaces, hypersurfaces, and surfaces in $\mathbb{R}^4$

Minimal surfaces in $\mathbb{R}^n$. If J denotes the natural complex structure on M, the 1-form

$$\tilde{\xi}(X) := \xi(X) - i\xi(JX), \quad X \in TM,$$

is $\mathbb{C}$-linear, with values in the complexified Clifford algebra $\tilde{\mathrm{Cl}}_n = \mathrm{Cl}_n \oplus i\,\mathrm{Cl}_n$; in general

$$F = \int \Re\mathrm{e}\,\tilde{\xi} = \int \Re\mathrm{e}(\tilde{f}(z)\,dz),$$

where z is a complex parameter of M and $\tilde{f}$ is a smooth function. Note that $\tilde{\xi}$ and $\tilde{f}$ in fact take their values in $\mathbb{C}^n := \mathbb{R}^n \oplus i\mathbb{R}^n \subset \tilde{\mathrm{Cl}}_n$.

Proposition 7.1. *The form $\tilde{\xi}$ is closed* (*and thus holomorphic*) *if and only if $\vec{H}=0$. In that case, we have*

$$F=\Re\mathrm{e}\int\tilde{f}(z)\,dz,$$

where $\tilde{f}$ is a holomorphic function.

Proof. We assume that (e_1,e_2) is a local orthonormal frame on M, positively oriented, such that $\nabla e_1=\nabla e_2=0$ at a point x_0. We thus have

$$d\tilde{\xi}(e_1,e_2)=\partial_{e_1}(\xi(e_2)+i\xi(e_1))-\partial_{e_2}(\xi(e_1)-i\xi(e_2)).$$

Noticing that, for $j,k\in\{1,2\}$,

$$\partial_{e_j}(\xi(e_k))=\partial_{e_j}\langle\langle e_k\cdot\varphi,\varphi\rangle\rangle=\langle\langle e_k\cdot\nabla_{e_j}\varphi,\varphi\rangle\rangle+\langle\langle e_k\cdot\varphi,\nabla_{e_j}\varphi\rangle\rangle=(\mathrm{id}+\tau)\langle\langle e_k\cdot\nabla_{e_j}\varphi,\varphi\rangle\rangle,$$

we obtain

$$d\tilde{\xi}(e_1,e_2)=i\,(\mathrm{id}+\tau)\langle\langle e_1\cdot\nabla_{e_1}\varphi+e_2\cdot\nabla_{e_2}\varphi,\,\varphi\rangle\rangle+(\mathrm{id}+\tau)\langle\langle e_2\cdot\nabla_{e_1}\varphi-e_1\cdot\nabla_{e_2}\varphi,\,\varphi\rangle\rangle.$$

The first term on the right-hand side is

$$i\,(\mathrm{id}+\tau)\langle\langle e_1\cdot\nabla_{e_1}\varphi+e_2\cdot\nabla_{e_2}\varphi,\,\varphi\rangle\rangle=i\,(\mathrm{id}+\tau)\langle\langle\vec{H}\cdot\varphi,\,\varphi\rangle\rangle=2i\,\langle\langle\vec{H}\cdot\varphi,\,\varphi\rangle\rangle,$$

since, by (8),

$$D\varphi:=e_1\cdot\nabla_{e_1}\varphi+e_2\cdot\nabla_{e_2}\varphi=\vec{H}\cdot\varphi,$$

and $\tau[\vec{H}]=[\vec{H}]$. The second term is

$$\begin{aligned}(\mathrm{id}+\tau)\langle\langle e_2\cdot\nabla_{e_1}\varphi-e_1\cdot\nabla_{e_2}\varphi,\,\varphi\rangle\rangle&=-(\mathrm{id}+\tau)\langle\langle e_1\cdot\nabla_{e_1}\varphi+e_2\cdot\nabla_{e_2}\varphi,\,e_1\cdot e_2\cdot\varphi\rangle\rangle\\&=-(\mathrm{id}+\tau)\langle\langle\vec{H}\cdot\varphi,\,e_1\cdot e_2\cdot\varphi\rangle\rangle\\&=0,\end{aligned}$$

using again that $D\varphi=\vec{H}\cdot\varphi$ and since $\tau([\vec{H}][e_1][e_2])=-[\vec{H}][e_1][e_2]$. We thus obtain the formula

$$d\tilde{\xi}(e_1,e_2)=2i\,\langle\langle\vec{H}\cdot\varphi,\,\varphi\rangle\rangle,$$

which may be written in the form

$$d\tilde{\xi}=-\mu^2\langle\langle\vec{H}\cdot\varphi,\,\varphi\rangle\rangle dz\wedge d\bar{z},\tag{17}$$

where μ is such that the metric is $\mu^2\,dz\,d\bar{z}$. This gives the first part of the lemma. Assuming that $\vec{H}=0$, the 1-form $\tilde{\xi}$ is closed, and the $\mathbb{C}^n$-valued function $\tilde{f}$ such that $\tilde{\xi}=\tilde{f}\,dz$ is holomorphic; the result follows. □

The aim now is to obtain explicit formulas in terms of holomorphic functions involving the components of the spinor field. We first note the following expression of $\tilde{f}$ in terms of the spinor field φ:

Lemma 7.2. *We have*

$$\tilde{f} = \mu\{\tau[\varphi]e_1^o[\varphi] - i\tau[\varphi]e_2^o[\varphi]\},$$

where the real function μ *is such that the metric is*

$$\mu^2(dx^2 + dy^2) \quad \text{in } z = x + iy,$$

$[\varphi]$ *represents the spinor field* φ *in a spinorial frame above* $((1/\mu)\partial_x, (1/\mu)\partial_y)$, *and* e_1^o, e_2^o *are the first two vectors of the canonical basis of* $\mathbb{R}^n \subset \mathrm{Cl}_n$.

Proof. We have

$$\tilde{f} = \tilde{\xi}(\partial_x) = \tau[\varphi][\partial_x][\varphi] - i\tau[\varphi][\partial_y][\varphi],$$

and the result follows since $[(1/\mu)\partial_x] = e_1^o$ and $[(1/\mu)\partial_y] = e_2^o$ in such a spinorial frame. □

Minimal surfaces in $\mathbb{R}^3$. Assuming that $n = 3$ and $H = 0$, we easily get by a computation using Lemma 7.2 that

$$F = \int \Re\mathrm{e}(\tilde{f}(z)\,dz) = \Re\mathrm{e}\left(\int \tilde{f}(z)\,dz\right),$$

where $\tilde{f} = \left(\frac{1}{2}if(1+g^2), \frac{1}{2}f(1-g^2), fg\right)$, with

$$f = 2\mu z_1^2, \quad g = -i\frac{\overline{z_2}}{z_1};$$

the complex functions z_1, z_2 are the components of φ in a spinorial frame above $((1/\mu)\partial_x, (1/\mu)\partial_y)$, and the functions f and g are holomorphic, since so are $\sqrt{\mu}z_1$ and $\sqrt{\mu}\overline{z_2}$ (this is a consequence of the Dirac equation $D\varphi = 0$, in $z = x + iy$). This is the classical Weierstrass representation of minimal surfaces in $\mathbb{R}^3$.

Minimal surfaces in $\mathbb{R}^4$. In the case of a surface in $\mathbb{R}^4$, we may also recover the explicit formulas of Konopelchenko and Landolfi [1999] expressing a general immersion in terms of 4 complex functions, which are solutions of first order PDEs; the functions are holomorphic if $\vec{H} = 0$. We do not include the calculations, since the general representation in Theorem 2 easily reduces to the spinor representation given in [Bayard et al. 2013] if $p = 2$ and $n = 4$ (see page 74), and the equivalence of this representation with the Konopelchenko–Landolfi representation is proved in [Romon and Roth 2013].

Remark. For surfaces in $\mathbb{R}^n$, $n \geq 5$, it is still possible to obtain an explicit representation in terms of the components of the spinor field which represents the surface, with holomorphic data if $\vec{H} = 0$, if the bundle E is assumed to be flat. We do not know if such a representation is possible without this additional assumption.

Hypersurfaces in $\mathbb{R}^n$**.** We set $p = n - 1$, and assume that M is a p-dimensional Riemannian manifold and E is the trivial line bundle on M, oriented by a unit section $\nu \in \Gamma(E)$. We moreover suppose that M is simply connected and that $h : TM \times TM \to \mathbb{R}$ is a given symmetric bilinear form. According to Theorem 2, an isometric immersion of M into $\mathbb{R}^{p+1}$ with normal bundle E and second fundamental form $B = h\nu$ is equivalent to a section φ of $\Gamma(U\Sigma)$ solution of the Killing equation (8). Note that $Q_E \simeq M$ and the double covering

$$\tilde{Q}_E \to Q_E$$

is trivial, since M is assumed to be simply connected. Fixing a section $\tilde{s}_E$ of $\tilde{Q}_E$ we get an injective map

$$\tilde{Q}_M \to \tilde{Q}_M \times_M \tilde{Q}_E =: \tilde{Q}, \quad \tilde{s}_M \mapsto (\tilde{s}_M, \tilde{s}_E).$$

Using

$$\mathrm{Cl}_p \simeq \mathrm{Cl}^0_{p+1} \subset \mathrm{Cl}_{p+1}$$

(induced by the Clifford map $\mathbb{R}^p \to \mathrm{Cl}_{p+1},\ X \mapsto X \cdot e_{p+1}$), we deduce a bundle isomorphism

$$\tilde{Q}_M \times_\rho \mathrm{Cl}_p \to \tilde{Q} \times_\rho \mathrm{Cl}^0_{p+1} \subset \Sigma, \quad \psi \mapsto \psi^*. \tag{18}$$

It satisfies the following properties: for all $X \in TM$ and $\psi \in \tilde{Q}_M \times_\rho \mathrm{Cl}_p$,

$$(X \cdot_M \psi)^* = X \cdot \nu \cdot \psi^* \quad \text{and} \quad \nabla_X(\psi^*) = (\nabla_X \psi)^*.$$

The section $\varphi \in \Gamma(U\Sigma)$ solution of (8) thus identifies to a section ψ of $\tilde{Q}_M \times_\rho \mathrm{Cl}_p$ solution of

$$\nabla_X \psi = -\frac{1}{2}\sum_{j=1}^{p} h(X, e_j)e_j \cdot_M \psi = -\frac{1}{2}T(X) \cdot_M \psi$$

for all $X \in TM$, where $T : TM \to TM$ is the symmetric operator associated to h. We deduce the following result:

Theorem 4. *Let $T : TM \to TM$ be a symmetric operator. The following two statements are equivalent*:

(1) *there exists an isometric immersion of M into $\mathbb{R}^{p+1}$ with shape operator T*;

(2) *there exists a normalized spinor field $\psi \in \Gamma(\tilde{Q}_M \times_\rho \mathrm{Cl}_p)$ solution of*

$$\nabla_X \psi = -\tfrac{1}{2}T(X) \cdot_M \psi \quad \textit{for all } X \in TM. \tag{19}$$

Here, a spinor field $\psi \in \Gamma(\tilde{Q}_M \times_\rho \mathrm{Cl}_p)$ is said to be normalized if it is represented in some frame $\tilde{s} \in \tilde{Q}_M$ by an element $[\psi] \in \mathrm{Cl}_p \simeq \mathrm{Cl}^0_{p+1}$ belonging to $\mathrm{Spin}(p+1)$.

We will see below explicit representation formulas in the cases of dimension 3 and 4.

Surfaces in $\mathbb{R}^3$. Since $\mathrm{Cl}_2 \simeq \Sigma_2$ we have

$$\tilde{Q}_M \times_\rho \mathrm{Cl}_2 \simeq \Sigma M,$$

and φ is equivalent to a normalized spinor field $\psi \in \Gamma(\Sigma M)$ solution of

$$\nabla_X \psi = -\tfrac{1}{2} T(X) \cdot_M \psi$$

for all $X \in TM$; this equation is also equivalent to the equation $D\psi = H\psi$. This is the result obtained by Friedrich [1998].

We now write the representation formula (9) using a special model for Cl_3, and indicate how to recover Friedrich's representation formula. We first consider the Clifford map

$$(x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto \begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix} \in \mathbb{H}(2),$$

where $x = -ix_3 + j(x_1 + ix_2)$, which identifies Cl_3 to the set

$$\left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, x, y \in \mathbb{H} \right\}$$

and $\mathbb{R}^3 \subset \mathrm{Cl}_3$ to the set of the imaginary quaternions; we also consider the ideal of Cl_3

$$\Sigma_3 = \left\{ \begin{pmatrix} y & 0 \\ 0 & 0 \end{pmatrix}, y \in \mathbb{H} \right\} \subset \mathrm{Cl}_3, \tag{20}$$

which is a model of the spin representation. Now φ, section of $U\Sigma = \tilde{Q} \times_\rho \mathrm{Spin}(3)$, is equivalent to a unit spinor field $\varphi' \in \Gamma(\tilde{Q} \times_\rho \Sigma_3)$ (obtained by projection) and a direct computation yields

$$\langle\langle X \cdot \varphi, \varphi \rangle\rangle = i\, \Im\mathrm{m}\langle X \cdot \varphi', \varphi' \rangle + j \langle X \cdot \varphi', \alpha(\varphi') \rangle \tag{21}$$

for all $X \in TM$, where the brackets $\langle \cdot\,, \cdot \rangle$ stand for the natural hermitian product on Σ_3 and $\alpha : \Sigma_3 \to \Sigma_3$ is the natural quaternionic structure. The representation formula given by the right-hand side term of (21) appears in [Friedrich 1998]. Finally, the identification (18) for the dimension $p = 2$

$$\tilde{Q}_M \times_\rho \mathrm{Cl}_2 \to \tilde{Q} \times_\rho \mathrm{Cl}_3^0 \subset \Sigma, \quad \psi \mapsto \psi^*$$

identifies $\varphi \in \Gamma(U\Sigma)$ to a unit spinor field $\psi \in \Gamma(\Sigma M)$, and it may be proved by a computation that

$$\langle\langle X \cdot \varphi, \varphi \rangle\rangle = i2\, \Re\mathrm{e}\langle X \cdot \psi^+, \psi^- \rangle + j(\langle X \cdot \psi^+, \alpha(\psi^+) \rangle - \langle X \cdot \psi^-, \alpha(\psi^-) \rangle),$$

where the brackets $\langle \cdot\,, \cdot \rangle$ stand here for the natural hermitian product on Σ_2 and $\alpha : \Sigma_2 \to \Sigma_2$ is the natural quaternionic structure; this is the explicit formula of the immersion in terms of ψ given in [Friedrich 1998].

Hypersurfaces in $\mathbb{R}^4$. Since $\mathrm{Cl}_3 \simeq \Sigma_3 \oplus \Sigma_3'$ where Σ_3 and Σ_3' are the two (nonequivalent) irreducible representations of Cl_3, we get two unit spinor fields $\psi_1 \in \Gamma(\Sigma M)$, $\psi_2 \in \Gamma(\Sigma' M)$ solutions of (19). Noting finally that there is a natural identification

$$i : \Sigma' M \to \Sigma M$$

satisfying

$$i(X \cdot \psi) = -X \cdot i(\psi)$$

for all $X \in TM$ and $\psi \in \Sigma' M$, the spinor fields ψ_1 and $i(\psi_2) \in \Gamma(\Sigma M)$ satisfy

$$\nabla_X \psi_1 = -\tfrac{1}{2} T(X) \cdot_M \psi_1 \quad \text{and} \quad \nabla_X i(\psi_2) = \tfrac{1}{2} T(X) \cdot_M i(\psi_2). \tag{22}$$

We thus recover a result of [Lawn and Roth 2010]: the immersion is equivalent to two spinor fields on the hypersurface which are solutions of (22). We may also obtain a new explicit representation formula. On one hand, we note that

$$\langle\langle X \cdot \varphi, \varphi \rangle\rangle = \begin{pmatrix} 0 & \bar{\xi}_1 x \xi_2 \\ \bar{\xi}_2 x \xi_1 & 0 \end{pmatrix} \tag{23}$$

in Cl_4^0, where $\varphi \in \Gamma(U\Sigma)$ and $X \in TM$ are respectively represented in Cl_4^0 by

$$\begin{pmatrix} \xi_1 & 0 \\ 0 & \xi_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix},$$

with $\xi_1, \xi_2 \in \mathbb{H}$ and $x \in \Im\mathrm{m}\,\mathbb{H}$. On the other hand, Σ_3 naturally identifies to $\mathbb{H}$ (see (20)) and the bilinear map

$$\Sigma_3 \times \Sigma_3 \to \mathbb{H}, \quad (\xi, \xi') \mapsto \bar{\xi}' \xi$$

induces a pairing

$$\langle\langle \cdot\,, \cdot \rangle\rangle_{\Sigma M} : \Sigma M \times \Sigma M \to \mathbb{H}$$

on $\Sigma M = \tilde{Q}_M \times_\rho \Sigma_3$. If

$$\psi = \psi_1 + \psi_2, \quad \psi_1 \in \Sigma M,\ \psi_2 \in \Sigma' M$$

is such that $\varphi = \psi^*$ (by (18), with $p = 3$), the spinor fields ψ_1 and $i(\psi_2) \in \Sigma M$ are respectively represented by ξ_1 and ξ_2, and we readily get

$$\langle\langle X \cdot_M i(\psi_2), \psi_1 \rangle\rangle_{\Sigma M} = \bar{\xi}_1 x \xi_2. \tag{24}$$

The identities (23) and (24) identify

$$\langle\langle X \cdot \varphi, \varphi \rangle\rangle \simeq \langle\langle X \cdot_M i(\psi_2), \psi_1 \rangle\rangle_{\Sigma M};$$

this gives an explicit representation of the immersion into $\mathbb{R}^4$ in terms of the two spinor fields ψ_1 and $i(\psi_2)$ of ΣM introduced in [Lawn and Roth 2010].

Surfaces in $\mathbb{R}^4$. For a surface in $\mathbb{R}^4$, Theorem 2 with $p = 2$ and $n = 4$ reduces to the result obtained in [Bayard et al. 2013], since the bundle Σ naturally identifies to the bundle $\Sigma M \otimes \Sigma E$ in that case (see the remark on page 55, observing that the representation of Spin(2) on Cl_2 by left multiplication is also the usual complex spin representation Σ_2). Note that we may similarly recover the main results in [Bayard 2013; Bayard and Patty 2015] concerning immersions in $\mathbb{R}^{3,1}$ and $\mathbb{R}^{2,2}$, if we consider in our constructions the Clifford algebras $\mathrm{Cl}_{3,1}$ and $\mathrm{Cl}_{2,2}$ instead of Cl_4.

For completeness, we also want to mention that in [Romon and Roth 2013], the authors give the explicit correspondence between the spinors used in [Bayard et al. 2013] for surfaces of $\mathbb{R}^4$ and a quaternionic representation which is a quaternionic reformulation of the representation obtained by Konopelchenko [2000] (see [Helein 2001] for this reformulation). The reader can refer to [Kamberov et al. 2002] for a detailed presentation of quaternionic-type representations of surfaces.

8. Spinorial representation of submanifolds in $\mathbb{S}^n$ and $\mathbb{H}^n$

We extend here Theorem 2 to the other space forms.

Submanifolds of $\mathbb{S}^n$. Let M be a Riemannian manifold of dimension p, and E be a bundle on M of rank $q = n - p$, with a fiber metric and a compatible connection; we assume that TM and E are spin, and consider

$$\Sigma := \tilde{Q} \times_\rho \mathrm{Cl}_{n+1},$$

where $\tilde{Q} = \tilde{Q}_M \times_M \tilde{Q}_E$ is the $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$ principal bundle given by the two spin structures and $\rho : \mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Aut}(\mathrm{Cl}_{n+1})$ is the representation obtained by the composition of the maps

$$\mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Spin}(n) \subset \mathrm{Spin}(n+1) \tag{25}$$

and

$$\mathrm{Spin}(n+1) \to \mathrm{Aut}(\mathrm{Cl}_{n+1}). \tag{26}$$

The maps in (25) correspond to the decompositions

$$\mathbb{R}^p \oplus \mathbb{R}^q =: \mathbb{R}^n \subset \mathbb{R}^n \oplus \mathbb{R}e_{n+1} =: \mathbb{R}^{n+1},$$

and in (26) the action of $\mathrm{Spin}(n+1)$ on Cl_{n+1} is the multiplication on the left. We also define

$$U\Sigma = \tilde{Q} \times_\rho \mathrm{Spin}(n+1) \subset \Sigma.$$

Let us denote by ν the element of the Clifford bundle $\tilde{Q} \times_{\mathrm{Ad}} \mathrm{Cl}_{n+1}$ such that its component in an arbitrary frame $\tilde{s} \in \tilde{Q}$ is the constant vector e_{n+1} (note that for all $g \in \mathrm{Spin}(p) \times \mathrm{Spin}(q) \subset \mathrm{Spin}(n) \subset \mathrm{Spin}(n+1)$, $\mathrm{Ad}(g)(e_{n+1}) = e_{n+1}$).

Theorem 5. *Let $B : TM \times TM \to E$ be a symmetric and bilinear map. The following two statements are equivalent*:

(1) *There exists an isometric immersion F of M into $\mathbb{S}^n$ with normal bundle E and second fundamental form B.*

(2) *There exists a spinor field $\varphi \in \Gamma(U\Sigma)$ satisfying*

$$\nabla_X \varphi = -\frac{1}{2}\sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot \varphi + \frac{1}{2} X \cdot \nu \cdot \varphi \quad \text{for all } X \in TM. \tag{27}$$

Moreover we have the representation formula

$$F = \langle\langle \nu \cdot \varphi, \varphi \rangle\rangle \in \mathbb{S}^n \subset \mathbb{R}^{n+1}, \tag{28}$$

where the brackets $\langle\langle \cdot, \cdot \rangle\rangle$ are defined as in (3)–(4).

Proof. We only prove that (2) implies (1), using the explicit formula (28). Setting $F = \langle\langle \nu \cdot \varphi, \varphi \rangle\rangle$, we have

$$F = [\varphi]^{-1} e_{n+1} [\varphi] = \mathrm{Ad}([\varphi]^{-1})(e_{n+1}),$$

where $[\varphi] \in \mathrm{Spin}(n+1)$ represents φ in some frame $\tilde{s} \in \tilde{Q}$ and $\mathrm{Ad} : \mathrm{Spin}(n+1) \to \mathrm{SO}(n+1)$ is the natural double covering; thus F belongs to $\mathbb{S}^n$. We will need the following:

Lemma 8.1. *If $\varphi \in \Gamma(U\Sigma)$ is a solution of* (27) *then $F = \langle\langle \nu \cdot \varphi, \varphi \rangle\rangle$ is such that, for all $X \in TM$,*

$$dF(X) = \langle\langle X \cdot \varphi, \varphi \rangle\rangle. \tag{29}$$

Proof. We first observe that $\nabla \nu = 0$: if α is the connection form on $\tilde{Q}$ and $\tilde{s} \in \Gamma(\tilde{Q})$ is a local frame, then $\nu = [\tilde{s}, e_{n+1}]$ and

$$\nabla_X \nu = [\tilde{s}, \partial_X e_{n+1} + \mathrm{Ad}_*(\alpha(\tilde{s}_*(X)))(e_{n+1})] = 0 \quad \text{for all } X \in TM,$$

since e_{n+1} is constant and α takes values in $\Lambda^2\mathbb{R}^n \subset \mathrm{Cl}_n$. Thus, for all $X \in TM$,

$$\begin{aligned} dF(X) &= \langle\langle \nu \cdot \nabla_X \varphi, \varphi \rangle\rangle + \langle\langle \nu \cdot \varphi, \nabla_X \varphi \rangle\rangle \\ &= (\mathrm{id} + \tau)\langle\langle \nu \cdot \nabla_X \varphi, \varphi \rangle\rangle \\ &= -\frac{1}{2}(\mathrm{id} + \tau)\sum_{j=1}^{p} \langle\langle \nu \cdot e_j \cdot B(X, e_j) \cdot \varphi, \varphi \rangle\rangle + \frac{1}{2}(\mathrm{id} + \tau)\langle\langle X \cdot \varphi, \varphi \rangle\rangle. \end{aligned}$$

But

$$\begin{aligned} \tau\langle\langle \nu \cdot e_j \cdot B(X, e_j) \cdot \varphi, \varphi \rangle\rangle &= \langle\langle \varphi, \nu \cdot e_j \cdot B(X, e_j) \cdot \varphi \rangle\rangle \\ &= \langle\langle B(X, e_j) \cdot e_j \cdot \nu \cdot \varphi, \varphi \rangle\rangle \\ &= -\langle\langle \nu \cdot e_j \cdot B(X, e_j) \cdot \varphi, \varphi \rangle\rangle \end{aligned}$$

since the three vectors $B(X, e_j)$, e_j and ν are mutually orthogonal, and

$$\tau\langle\langle X\cdot\varphi,\varphi\rangle\rangle=\langle\langle\varphi, X\cdot\varphi\rangle\rangle=\langle\langle X\cdot\varphi,\varphi\rangle\rangle.$$

Thus (29) follows. □

By the lemma and the properties of the Clifford product, F is an isometric immersion, and the map

$$E\to T\mathbb{S}^n,\quad X\in E_m\mapsto(F(m),\langle\langle X\cdot\varphi,\varphi\rangle\rangle)$$

identifies E with the normal bundle of $F(M)$ into $\mathbb{S}^n$; it moreover identifies the connection on E with the normal connection of $F(M)$ in $\mathbb{S}^n$ and B with the second fundamental form. We omit the proof since it is very similar to the proof of Proposition 4.2. □

Remark. Taking the trace of (27) we get

$$D\varphi=\tfrac{1}{2}p(\vec{H}-\nu)\cdot\varphi, \tag{30}$$

where $\vec{H}=(1/p)\sum_{j=1}^{p}B(e_j,e_j)$ is the mean curvature vector of M in $\mathbb{S}^n$.

Remark. We may also obtain a proof using spinors of the fundamental theorem of submanifold theory in $\mathbb{S}^n$, showing, as in Section 5, that the equations of Gauss, Codazzi and Ricci in a space of constant sectional curvature 1 are exactly the integrability conditions of (27).

We finally show how to recover the spinorial characterization of a surface in $\mathbb{S}^3$ given by Morel in [2005] and Taĭmanov in [2004]. In the model $\mathrm{Cl}_4\simeq\mathbb{H}(2)$ we have

$$\varphi=\begin{pmatrix}[\varphi^+] & 0\\ 0 & [\varphi^-]\end{pmatrix},\quad F=\begin{pmatrix}0 & \overline{[\varphi^+]}[\nu][\varphi^-]\\ -\overline{[\varphi^-]}[\nu][\varphi^+] & 0\end{pmatrix}$$

and

$$\xi(X)=\begin{pmatrix}0 & \overline{[\varphi^+]}[X][\varphi^-]\\ -\overline{[\varphi^-]}[X][\varphi^+] & 0\end{pmatrix},$$

where $[\varphi^+]$, $[\varphi^-]$, $[\nu]$ and $[X]\in\mathbb{H}$ represent φ^+, φ^-, ν and X in some spinor frame adapted to the immersion in $\mathbb{S}^3$; thus Lemma 8.1 gives

$$F\simeq\overline{[\varphi^+]}[\nu][\varphi^-]\quad\text{and}\quad dF(X)\simeq\overline{[\varphi^+]}[X][\varphi^-].$$

If $[\varphi^+]$ is given, this system has a solution $[\varphi^-]$, unique up to the multiplication by $\mathbb{S}^3$ on the right. The spinor field φ is thus essentially determined by its component φ^+, which may be identified with a spinor field $\psi\in\Gamma(\Sigma M)$ solution of

$$D\psi=H\psi-i\overline{\psi},\quad|\psi|=1;$$

details are given in [Bayard et al. 2013]. This is the spinor characterization of an immersion in $\mathbb{S}^3$ given in [Morel 2005; Taĭmanov 2004].

Submanifolds of $\mathbb{H}^n$***.*** We now consider the n-dimensional hyperbolic space $\mathbb{H}^n$ as a hypersurface of the Minkowski space $\mathbb{R}^{n,1}$. Since the constructions of the paper may also be carried out in a linear space with a semi-Riemannian metric, we obtain a spinor representation of a submanifold in $\mathbb{H}^n$ exactly as we did for a submanifold in $\mathbb{S}^n$. We thus only state the results here, and refer to the previous section for the proofs. Let M be a Riemannian manifold of dimension p, and E be a bundle on M of rank $q = n - p$, with a Riemannian fiber metric and a compatible connection; we assume that TM and E are spin, and consider

$$\Sigma := \tilde{Q} \times_\rho \mathrm{Cl}_{n,1},$$

where $\tilde{Q} = \tilde{Q}_M \times_M \tilde{Q}_E$ is the $\mathrm{Spin}(p) \times \mathrm{Spin}(q)$ principal bundle given by the two spin structures and $\rho : \mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Aut}(\mathrm{Cl}_{n,1})$ is the representation obtained by the composition of the maps

$$\mathrm{Spin}(p) \times \mathrm{Spin}(q) \to \mathrm{Spin}(n) \subset \mathrm{Spin}(n, 1) \tag{31}$$

and

$$\mathrm{Spin}(n, 1) \to \mathrm{Aut}(\mathrm{Cl}_{n,1}). \tag{32}$$

The maps in (31) correspond to the decompositions

$$\mathbb{R}^p \oplus \mathbb{R}^q =: \mathbb{R}^n \subset \mathbb{R}^n \oplus \mathbb{R}e_{n+1} =: \mathbb{R}^{n,1},$$

and in (32) the action of $\mathrm{Spin}(n, 1)$ on $\mathrm{Cl}_{n,1}$ is the multiplication on the left; here e_{n+1} is a vector with negative norm -1. We also define

$$U\Sigma = \tilde{Q} \times_\rho \mathrm{Spin}(n, 1) \subset \Sigma.$$

Let us denote by ν the element of the Clifford bundle $\tilde{Q} \times_{\mathrm{Ad}} \mathrm{Cl}_{n,1}$ such that its component in an arbitrary frame $\tilde{s} \in \tilde{Q}$ is the constant vector e_{n+1}.

Theorem 6. *Let $B : TM \times TM \to E$ be a symmetric and bilinear map. The following two statements are equivalent*:

1. *There exists an isometric immersion F of M into $\mathbb{H}^n$ with normal bundle E and second fundamental form B.*
2. *There exists a spinor field $\varphi \in \Gamma(U\Sigma)$ satisfying*

$$\nabla_X \varphi = -\frac{1}{2} \sum_{j=1}^{p} e_j \cdot B(X, e_j) \cdot \varphi - \frac{1}{2} X \cdot \nu \cdot \varphi \quad \textit{for all } X \in TM. \tag{33}$$

Moreover we have the representation formula

$$F = \langle\langle \nu \cdot \varphi, \varphi \rangle\rangle \in \mathbb{H}^n \subset \mathbb{R}^{n,1}, \tag{34}$$

where the brackets $\langle\langle \cdot , \cdot \rangle\rangle$ are defined as in (3)-(4).

We may also recover the spinor characterization of an immersion of a surface in $\mathbb{H}^3$ given by Morel [2005]: if M is a surface and (e_1, e_2) is an orthonormal basis of TM, setting $\vec{H}_{\mathbb{H}^3} := \frac{1}{2}(B(e_1, e_1) + B(e_2, e_2))$ we see that (33) is equivalent to

$$D\varphi = (\vec{H}_{\mathbb{H}^3} + \nu) \cdot \varphi,$$

where φ is a spinor field which is represented in a frame $\tilde{s} \in \tilde{Q}$ by $[\varphi]$ belonging to Spin(3, 1). This is exactly the spinor representation of an immersion in $\mathbb{H}^3$ as described in [Bayard 2013] Section 5, where it is moreover proved that it is equivalent to the spinor characterization given in [Morel 2005].

Remark. The Weierstrass-type representation of the flat surfaces in $\mathbb{H}^3$ by J.A. Gálvez, A. Martínez and F. Milán [Gálvez et al. 2000] may be recovered from Theorem 6, following the lines of Section 6.4 in [Bayard 2013]: the explicit representation is (34), whereas equation (33) gives rise to holomorphic data, when written in conformal coordinates induced by the Gauss map; we refer to these papers for details.

Acknowledgments

P. Bayard was supported by the project PAPIIT IA105116.

References

[Ammann et al. 2013] B. Ammann, A. Moroianu, and S. Moroianu, "The Cauchy problems for Einstein metrics and parallel spinors", *Comm. Math. Phys.* **320**:1 (2013), 173–198. MR Zbl

[Bär 1998] C. Bär, "Extrinsic bounds for eigenvalues of the Dirac operator", *Ann. Global Anal. Geom.* **16**:6 (1998), 573–596. MR Zbl

[Bär et al. 2005] C. Bär, P. Gauduchon, and A. Moroianu, "Generalized cylinders in semi-Riemannian and spin geometry", *Math. Z.* **249**:3 (2005), 545–580. MR Zbl

[Bayard 2013] P. Bayard, "On the spinorial representation of spacelike surfaces into 4-dimensional Minkowski space", *J. Geom. Phys.* **74** (2013), 289–313. MR Zbl

[Bayard and Patty 2015] P. Bayard and V. Patty, "Spinor representation of Lorentzian surfaces in $\mathbb{R}^{2,2}$", *J. Geom. Phys.* **95** (2015), 74–95. MR Zbl

[Bayard et al. 2013] P. Bayard, M.-A. Lawn, and J. Roth, "Spinorial representation of surfaces into 4-dimensional space forms", *Ann. Global Anal. Geom.* **44**:4 (2013), 433–453. MR Zbl

[Chiossi and Salamon 2002] S. Chiossi and S. Salamon, "The intrinsic torsion of SU(3) and G_2 structures", pp. 115–133 in *Differential geometry* (Valencia, 2001), edited by O. Gil-Medrano and V. Miquel, World Sci. Publ., River Edge, NJ, 2002. MR Zbl

[Friedrich 1998] T. Friedrich, "On the spinor representation of surfaces in Euclidean 3-space", *J. Geom. Phys.* **28**:1-2 (1998), 143–157. MR Zbl

[Gálvez et al. 2000] J. A. Gálvez, A. Martínez, and F. Milán, "Flat surfaces in the hyperbolic 3-space", *Math. Ann.* **316**:3 (2000), 419–435. MR Zbl

[Helein 2001] F. Helein, "On Konopelchenko's representation formula for surfaces in 4 dimensions", preprint, 2001. arXiv

[Kamberov et al. 2002] G. Kamberov, P. Norman, F. Pedit, and U. Pinkall, *Quaternions, spinors, and surfaces*, Contemporary Mathematics **299**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Kenmotsu 1979] K. Kenmotsu, "Weierstrass formula for surfaces of prescribed mean curvature", *Math. Ann.* **245**:2 (1979), 89–99. MR Zbl

[Konopelchenko 1996] B. G. Konopelchenko, "Induced surfaces and their integrable dynamics", *Stud. Appl. Math.* **96**:1 (1996), 9–51. MR Zbl

[Konopelchenko 2000] B. G. Konopelchenko, "Weierstrass representations for surfaces in 4D spaces and their integrable deformations via DS hierarchy", *Ann. Global Anal. Geom.* **18**:1 (2000), 61–74. MR Zbl

[Konopelchenko and Landolfi 1999] B. G. Konopelchenko and G. Landolfi, "Generalized Weierstrass representation for surfaces in multi-dimensional Riemann spaces", *J. Geom. Phys.* **29**:4 (1999), 319–333. MR Zbl

[Konopelchenko and Landolfi 2000] B. G. Konopelchenko and G. Landolfi, "Induced surfaces and their integrable dynamics, II: Generalized Weierstrass representations in 4D spaces and deformations via DS hierarchy", *Stud. Appl. Math.* **104**:2 (2000), 129–169. MR Zbl

[Konopelchenko and Taĭmanov 1996] B. G. Konopelchenko and I. A. Taĭmanov, "Constant mean curvature surfaces via an integrable dynamical system", *J. Phys. A* **29**:6 (1996), 1261–1265. MR Zbl

[Kusner and Schmitt 1996] R. Kusner and N. Schmitt, "The spinor representation of surfaces in space", preprint, 1996. arXiv

[Lawn 2008] M.-A. Lawn, "Immersions of Lorentzian surfaces in $\mathbb{R}^{2,1}$", *J. Geom. Phys.* **58**:6 (2008), 683–700. MR Zbl

[Lawn and Roth 2010] M.-A. Lawn and J. Roth, "Isometric immersions of hypersurfaces in 4-dimensional manifolds via spinors", *Differential Geom. Appl.* **28**:2 (2010), 205–219. MR Zbl

[Lawn and Roth 2011] M.-A. Lawn and J. Roth, "Spinorial characterizations of surfaces into 3-dimensional pseudo-Riemannian space forms", *Math. Phys. Anal. Geom.* **14**:3 (2011), 185–195. MR Zbl

[Lounesto 2001] P. Lounesto, *Clifford algebras and spinors*, 2nd ed., London Mathematical Society Lecture Note Series **286**, Cambridge Univ. Press, 2001. MR Zbl

[Morel 2005] B. Morel, "Surfaces in $\mathbb{S}^3$ and $\mathbb{H}^3$ via spinors", pp. 131–144 in *Actes du Séminaire de théorie spectrale et géométrie* (Grenoble, 2004–2005), vol. 23, Institut Fourier, Saint-Martin-d'Hères, 2005. MR Zbl

[Moroianu and Semmelmann 2014] A. Moroianu and U. Semmelmann, "Generalized Killing spinors and Lagrangian graphs", *Differential Geom. Appl.* **37** (2014), 141–151. MR Zbl

[Nakad 2011] R. Nakad, "The energy-momentum tensor on Spin^c manifolds", *Int. J. Geom. Methods Mod. Phys.* **8**:2 (2011), 345–365. MR Zbl

[Nakad and Roth 2012] R. Nakad and J. Roth, "Hypersurfaces of Spin^c manifolds and Lawson type correspondence", *Ann. Global Anal. Geom.* **42**:3 (2012), 421–442. MR Zbl

[Romon and Roth 2013] P. Romon and J. Roth, "The spinor representation formula in 3 and 4 dimensions", pp. 261–282 in *Pure and applied differential geometry: PADGE 2012*, edited by J. Van der Veken et al., Shaker Verlag, Aachen, 2013. Zbl arXiv

[Roth 2010] J. Roth, "Spinorial characterizations of surfaces into three-dimensional homogeneous manifolds", *J. Geom. Phys.* **60**:6-8 (2010), 1045–1061. MR Zbl

[Taĭmanov 1997a] I. A. Taĭmanov, "Modified Novikov–Veselov equation and differential geometry of surfaces", pp. 133–151 in *Solitons, geometry, and topology: on the crossroad*, edited by V. M. Buchstaber and S. P. Novikov, Amer. Math. Soc. Transl. Ser. 2 **179**, American Mathematical Society, Providence, RI, 1997. MR Zbl

[Taĭmanov 1997b] I. A. Taĭmanov, "Surfaces of revolution in terms of solitons", *Ann. Global Anal. Geom.* **15**:5 (1997), 419–435. MR Zbl

[Taĭmanov 2004] I. A. Taĭmanov, "Dirac operators and conformal invariants of tori in three-dimensional space", *Proc. Steklov Inst. Math.* **244** (2004), 233–263. MR Zbl arXiv

Received May 23, 2016. Revised February 7, 2017.

PIERRE BAYARD
FACULTAD DE CIENCIAS
UNAM
AV. UNIVERSIDAD 3000, CIRCUITO EXTERIOR S/N
CIUDAD UNIVERSITARIA
04510 CIUDAD DE MÉXICO,
MEXICO
bayard@ciencias.unam.mx

MARIE-AMÉLIE LAWN
DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
LONDON
SW7 2AZ
UNITED KINGDOM
m.lawn@imperial.ac.uk

JULIEN ROTH
LABORATOIRE D'ANALYSE ET DE MATHÉMATIQUES APPLIQUÉES
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE
77454 MARNE-LA-VALLÉE CEDEX 2
FRANCE
julien.roth@u-pem.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.81

COMPACT COMPOSITION OPERATORS WITH NONLINEAR SYMBOLS ON THE H^2 SPACE OF DIRICHLET SERIES

FRÉDÉRIC BAYART AND OLE FREDRIK BREVIG

We investigate compactness of composition operators on the Hardy space of Dirichlet series induced by a map $\varphi(s) = c_0 s + \varphi_0(s)$, where φ_0 is a Dirichlet polynomial. Our results depend heavily on the characteristic c_0 of φ and, when $c_0 = 0$, on both the degree of φ_0 and its local behavior near a boundary point. We also study the approximation numbers for some of these operators. Our methods involve geometric estimates of Carleson measures and tools from differential geometry.

1. Introduction

A theorem of Gordon and Hedenmalm [1999] describes the bounded composition operators on the Hilbert space $\mathscr{H}^2$ of Dirichlet series,

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s},$$

with square summable coefficients endowed with the norm $\|f\|_{\mathscr{H}^2}^2 := \sum_{n=1}^{\infty} |a_n|^2$. We let $\mathbb{C}_\theta$ denote the half-plane of complex numbers $s = \sigma + it$ with $\sigma > \theta$. The Dirichlet series in $\mathscr{H}^2$ represent analytic functions in $\mathbb{C}_{1/2}$ and a mapping φ of $\mathbb{C}_{1/2}$ into itself defines a function $\mathscr{C}_\varphi(f) := f \circ \varphi$ on $\mathbb{C}_{1/2}$, if $f \in \mathscr{H}^2$. The operator $\mathscr{C}_\varphi : \mathscr{H}^2 \to \mathscr{H}^2$ is well defined and bounded if and only if φ is a member of the following class:

Definition. The *Gordon–Hedenmalm class*, denoted $\mathscr{G}$, is the set of functions $\varphi : \mathbb{C}_{1/2} \to \mathbb{C}_{1/2}$ of the form

$$\varphi(s) = c_0 s + \sum_{n=1}^{\infty} c_n n^{-s} =: c_0 s + \varphi_0(s), \tag{1}$$

Brevig is supported by Grant 227768 of the Research Council of Norway.
MSC2010: primary 47B33; secondary 30B50, 30H10.
Keywords: Dirichlet series, composition operators, approximation numbers.

where c_0 is a nonnegative integer called the *characteristic* of φ, the Dirichlet series φ_0 converges uniformly in $\mathbb{C}_\varepsilon$ ($\varepsilon > 0$) and has the following mapping properties:

(a) If $c_0 = 0$, then $\varphi_0(\mathbb{C}_0) \subset \mathbb{C}_{1/2}$.

(b) If $c_0 \geq 1$, then either $\varphi_0 \equiv 0$ or $\varphi_0(\mathbb{C}_0) \subset \mathbb{C}_0$.

Since the paper of Gordon and Hedenmalm, several authors have studied the properties of composition operators acting on $\mathscr{H}^2$ or on similar spaces of Dirichlet series (see for instance [Bayart 2002; 2003; Finet and Queffélec 2004; Finet et al. 2004; Queffélec and Seip 2015a]). In the present work, we are interested in the study of the compactness of $\mathscr{C}_\varphi$ when φ is a polynomial symbol, say

$$\varphi(s) = c_0 s + c_1 + \sum_{n=2}^{N} c_n n^{-s}, \tag{2}$$

and we implicitly assume that $\varphi \in \mathscr{G}$. The symbol φ is said to have *unrestricted range* if

$$\inf_{s \in \mathbb{C}_0} \operatorname{Re}(\varphi(s)) = \begin{cases} \frac{1}{2} & \text{if } c_0 = 0, \\ 0 & \text{if } c_0 \geq 1. \end{cases}$$

Correspondingly, if $\varphi(\mathbb{C}_0)$ is strictly contained in any smaller half-plane, we say that $\mathscr{C}_\varphi$ has *restricted range*. It is well known that the composition operator $\mathscr{C}_\varphi$ is compact when φ has restricted range [Bayart 2002, Theorem 21]. In what follows, we will assume that φ has unrestricted range.

Definition. A set of integers $\Lambda \subseteq \mathbb{N} - \{1\}$ is called $\mathbb{Q}$-*independent* if the set $\{\log n : n \in \Lambda\}$ is linearly independent over $\mathbb{Q}$.

Symbols of the form (2) have been extensively studied in the *linear case*,

$$\varphi(s) = c_0 s + c_1 + \sum_{j=1}^{d} c_{q_j} q_j^{-s}, \tag{3}$$

where the set $\{q_j\}$ is $\mathbb{Q}$-independent and $c_{q_j} \neq 0$. When $c_0 \geq 1$, it is proven in [Bayart 2003] that the operator $\mathscr{C}_\varphi$ is compact if and only if φ has restricted range. Our first result extends this to the case of an arbitrary polynomial:

Theorem 1. *Let φ be a Dirichlet polynomial of the form* (2) *with* $c_0 \geq 1$. *Then $\mathscr{C}_\varphi$ is compact if and only if φ has restricted range.*

As is to be expected when investigating composition operators on $\mathscr{H}^2$, the symbols with $c_0 = 0$ are more difficult to handle and require different techniques. In this case, it is proven independently in [Bayart 2003] and [Finet et al. 2004] that composition operators induced by linear symbols (3) with $c_0 = 0$ are compact if and only if φ has restricted range or $d \geq 2$.

The main effort of this paper is dedicated to extending this result to general polynomials. We rely crucially on a geometric description of such compact composition operators found in [Queffélec and Seip 2015a] (see Lemma 5 below). Our second result is:

Theorem 2. *Suppose that $\{q_j\}_{j=1}^d$ are $\mathbb{Q}$-independent and that*

$$\varphi(s) = \sum_{j=1}^{d} P_j(q_j^{-s})$$

is in $\mathcal{G}$, and that the polynomials P_j are nonconstant. Then $\mathcal{C}_\varphi$ is compact if and only if φ has restricted range or $d \geq 2$.

Theorem 2 is truly a nonlinear extension of the results for linear symbols, however it fails to handle the relatively simple cases

$$\varphi_1(s) = \tfrac{9}{2} - 2^{-s} - 3^{-s} - 2\cdot 6^{-s} \quad \text{and} \quad \varphi_2(s) = \tfrac{13}{2} - 4\cdot 2^{-s} - 4\cdot 3^{-s} + 2\cdot 6^{-s}, \tag{4}$$

where "mixed terms" are present. However, the compactness of the associated operators can be decided by our main result. Before this result can be stated, we need to introduce some additional definitions.

Definition. Let $\Lambda \subseteq \mathbb{N} - \{1\}$. We let the *complex dimension* of Λ, denoted $\mathfrak{D}(\Lambda)$, be the infimum of $\operatorname{card}(\Lambda_0)$ where $\Lambda_0 \subset \mathbb{N} - \{1\}$ is $\mathbb{Q}$-independent and multiplicatively generates Λ.

At this point, we should mention that the set Λ_0 attaining such an infimum is not necessarily unique. This is easily seen by considering $\Lambda = \{2^2\cdot 3^2, 2^4\cdot 3^2, 2^2\cdot 3^4, 2^4\cdot 3^4\}$, where Λ_0 can be chosen as any of the following sets:

$$\{2, 3\}, \qquad \{2^2, 3\}, \qquad \{2, 3^2\}, \qquad \{2^2, 3^2\}, \qquad \{2^2\cdot 3, 3\}, \qquad \{2, 2\cdot 3^2\}.$$

Now, we will rewrite (2) as

$$\varphi(s) = c_1 + \sum_{n\in\Lambda} c_n n^{-s} \tag{5}$$

with $c_n \neq 0$ for every $n \in \Lambda$. We pick some $\Lambda_0 = \{q_1, q_2, \dots, q_d\}$ where $d = \mathfrak{D}(\Lambda)$. Since Λ_0 generates Λ, any $n \in \Lambda$ can be written uniquely as a product of elements in Λ_0,

$$n = \prod_{j=1}^{d} q_j^{\alpha_j}.$$

This associates to n the d-dimensional multi-index $\alpha(n)$. Clearly, $\alpha(n)$ depends on the choice of Λ_0 as the example considered above illustrates.

Definition. The *degree of φ with respect to* Λ_0 is defined by

$$\deg(\varphi, \Lambda_0) = \sup\{|\alpha(n)| = \alpha_1 + \alpha_2 + \cdots + \alpha_d : n \in \Lambda\}.$$

Among the different Λ_0 which generate Λ and with $\operatorname{card}(\Lambda_0) = \mathfrak{D}(\Lambda)$, we choose an optimal Λ_0 in the sense that it minimizes $\deg(\varphi, \Lambda_0)$. The *degree* of φ is then equal to the value of $\deg(\varphi, \Lambda_0)$ where Λ_0 is optimal in the previous sense.

It is clear that there can be more than one optimal Λ_0, as the example considered above again demonstrates, where the three final possibilities all have $\deg(\varphi, \Lambda_0) = 4$ if φ is given by (5).

Remark. For maps of the form (3) as considered before, the complex dimension is equal to d and the degree is equal to 1, which justifies our terminology "linear case".

The study of the Hardy space of Dirichlet series $\mathcal{H}^2$ is intimately related to function theory on polydiscs. In our concerns, the main tool will be the so-called Bohr lift. Indeed, consider an optimal Λ_0 and use the substitution $q_j^{-s} \mapsto z_j$. To simplify the expressions in what follows, we will also subtract $\frac{1}{2}$. Hence we obtain a polynomial in d variables with the same degree as φ,

$$\Phi(z) = \left(c_1 - \tfrac{1}{2}\right) + \sum_{n \in \Lambda} c_n z^{\alpha(n)}. \tag{6}$$

The polynomial Φ will be called an *optimal Bohr lift* of φ. Using Kronecker's theorem (see for instance [Hardy and Wright 1979, Chapter 13]), the $\mathbb{Q}$-independence of Λ_0 implies that Φ maps $\mathbb{D}^d$ into $\mathbb{C}_0$. The polynomial Φ induces a map, denoted by ϕ, on $\mathbb{R}^d$ defined by

$$\phi(\theta_1, \theta_2, \ldots, \theta_d) = \Phi(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_d}).$$

Remark. We will sometimes need to define the Bohr lift when the map $\varphi(s) = \sum_{n \geq 1} c_n n^{-s}$ is not a Dirichlet polynomial. It is then defined as

$$\Phi(z) = \left(c_1 - \tfrac{1}{2}\right) + \sum_{n \geq 2} c_n z^{\alpha(n)}$$

where we use the substitution $p_j^{-s} \mapsto z_j$. If we assume that $\varphi \in \mathcal{G}$, its Bohr lift Φ is now well defined on $\mathbb{D}^\infty \cap c_0$, and Kronecker's theorem shows that this set is mapped by Φ into $\mathbb{C}_0$.

Let us come back to a polynomial $\varphi \in \mathcal{G}$. If we assume that φ has unrestricted range, there exists at least one point $w \in \mathbb{T}^d$ so that $\operatorname{Re} \Phi(w) = 0$, by the compactness of $\mathbb{T}^d$. Let $w = (e^{i\vartheta_1}, e^{i\vartheta_2}, \ldots, e^{i\vartheta_d})$. Then $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_d)$ has to be a critical point of $\operatorname{Re} \phi$ since this last map admits a minimum at ϑ. Moreover, the mapping properties of φ implies that the Hessian matrix of $\operatorname{Re} \phi$ at ϑ should be nonnegative.

Definition. We define the *boundary index of* Φ *at* w as the nonnegative integer $J(\Phi, w)$ such that the signature of the Hessian matrix of $\operatorname{Re}\phi$ at ϑ is equal to $(J(\Phi, w), 0)$.

With these definitions at hand, we are able to state our main theorem which shows that, when there are mixed terms, the complex dimension does not give enough information and that we need a more careful study of φ.

Theorem 3. *Let* $\varphi(s) = c_1 + \sum_{n\geq 2} c_n n^{-s}$ *be a Dirichlet polynomial in* $\mathscr{G}$ *with unrestricted range. Suppose that its complex dimension* d *is greater than or equal to* 2, *and let* Φ *be a minimal Bohr lift of* φ. *Assume that*

- *either the degree of* φ *is equal to* 1 *or* 2,
- *or the degree of* φ *is at least* 3 *and for any* $w \in \mathbb{T}^d$, *either* $\operatorname{Re}\Phi(w) > 0$ *or* $\operatorname{Re}\Phi(w) = 0$ *and* $J(\Phi, w) \geq 2$.

Then $\mathscr{C}_\varphi$ *is compact on* $\mathscr{H}^2$. *Moreover*, *the result is optimal in the following sense*:

- *If the complex dimension of* φ *is equal to* 1, *then* $\mathscr{C}_\varphi$ *is never compact.*
- *There exist polynomials* $\varphi \in \mathscr{G}$ *of arbitrary complex dimension and of arbitrary degree greater than or equal to* 3 *such that* $\mathscr{C}_\varphi$ *is not compact.*

At this point we should mention that Theorem 3 does not encompass Theorem 2, and we will return to this point later (see Section 7). However, Theorem 3 allows us to conclude that for the Dirichlet polynomials φ given by (4), which have complex dimension and degree equal to 2, the induced composition operators are compact.

We are also interested in the degree of compactness of our operators, which may be estimated using their approximation numbers.

Definition. Let H be a Hilbert space and let $T \in \mathfrak{L}(H)$. The n-th approximation number of T, denoted $a_n(T)$, is the distance of T to the operators of rank $< n$.

The study of the behavior of $a_n(\mathscr{C}_\varphi)$ when $\varphi \in \mathscr{G}$ is a linear symbol (3) has been done in [Queffélec and Seip 2015a]. In particular, it is shown there that

$$\left(\frac{1}{n}\right)^{(d-1)/2} \ll a_n(\mathscr{C}_\varphi) \ll \left(\frac{\log n}{n}\right)^{(d-1)/2},$$

where d is the complex dimension of φ. We will extend this result to a general context. To keep this introduction sufficiently short, we refer to Section 8 for our statement, and give only one striking consequence of it: we may distinguish the Schatten classes of linear operators on $\mathscr{H}^2$ using composition operators induced by polynomial symbols. By definition, a compact linear operator T belongs to the Schatten class S_p, for $0 < p < \infty$, if

$$\|T\|_p^p := \operatorname{Tr}(|T|^p) = \sum_{n=1}^{\infty} a_n(T)^p < \infty.$$

Corollary 4. *Let* $0 < p < q$. *There exists a Dirichlet polynomial* $\varphi \in \mathscr{G}$ *such that* $\mathscr{C}_\varphi \in S_q \setminus S_p$.

Let us end this introduction by mentioning that the composition operators induced by the maps φ_1 and φ_2 have different degrees of compactness. Indeed, we will show that

$$\left(\frac{1}{n}\right)^{1/2} \ll a_n(\mathscr{C}_{\varphi_1}) \ll \left(\frac{\log n}{n}\right)^{1/2} \quad \text{and} \quad \left(\frac{1}{n}\right)^{1/3} \ll a_n(\mathscr{C}_{\varphi_2}) \ll \left(\frac{\log n}{n}\right)^{1/3}.$$

Organization. The remainder of this paper is divided into seven sections.

- Section 2 contains the proof of Theorem 1. The content of this section is independent from that of the following sections.
- In Section 3 we introduce some necessary tools and results needed for the proof of Theorem 2 and Theorem 3.
- Section 4 is devoted to the proof of Theorem 2.
- Section 5 contains the proof of Theorem 3.
- In Section 6 we prove Lemma 12, which is the most technical part of Theorem 3.
- In Section 7 we discuss the case $\deg(\varphi) \geq 3$ and $J(\Phi, w) = 0$, its connection to Theorem 2 and some related examples.
- Finally, in Section 8, we discuss the decay of the sequence of approximation numbers for some of our operators.

Notation. The notation $f(\varepsilon) \ll g(\varepsilon)$ will mean that $f(\varepsilon) \leq Cg(\varepsilon)$ for some constant C which does not depend on ε. We will sometimes write $f(\varepsilon) \ll_a g(\varepsilon)$ to emphasize that C depends on a. As usual, we let $\{p_j\}$ denote the sequence of prime numbers written in increasing order. We let $\boldsymbol{m}_d$ denote the normalized Lebesgue measure on $\mathbb{T}^d$. This measure is invariant under rotations. If we do not have a priori knowledge of the complex dimension d, we will often call this measure $\boldsymbol{m}_\infty$. For a point $z = e^{i\theta}$ on the unit circle $\mathbb{T}$, we will always assume that $\theta \in (-\pi, \pi]$. Finally, $\mathbf{0}$ will denote the point $(0, \dots, 0) \in \mathbb{C}^d$, and $\mathbf{1}$ will similarly denote the point $(1, \dots, 1)$.

2. Proof of Theorem 1

Let $\varphi(s) = c_0 s + c_1 + \sum_{n=2}^N c_n n^{-s} \in \mathscr{G}$ such that $c_0 \geq 1$. We already know that if φ has restricted range, then $\mathscr{C}_\varphi$ is compact. Let us therefore assume that $\mathscr{C}_\varphi$ is compact and also assume that φ has unrestricted range, to argue by contradiction.

By [Bayart 2003, Theorem 3], we know that

$$\frac{\operatorname{Re}\varphi(s)}{\operatorname{Re}(s)} \xrightarrow{\operatorname{Re}(s)\to 0} +\infty. \tag{7}$$

Now, since φ has unrestricted range there exists a sequence $\{s_k = \sigma_k + it_k\}_{k\geq 1}$ in $\mathbb{C}_0$ such that $\operatorname{Re}\varphi(s_k) \to 0$. It is well known that this forces that $\sigma_k \to 0$ (see [Bayart 2003]). Then

$$\operatorname{Re}\varphi(s_k) = c_0\sigma_k + \operatorname{Re}(c_1) + \sum_{n=2}^{N} n^{-\sigma_k}(\operatorname{Re}(c_n)\cos(t_k\log(n)) + \operatorname{Im}(c_n)\sin(t_k\log(n))).$$

By successive extraction of subsequences, we may assume that there exist real numbers a_n and b_n so that for $2 \leq n \leq N$ we have, as $k \to \infty$,

$$\cos(t_k\log(n)) \to a_n \quad \text{and} \quad \sin(t_k\log(n)) \to b_n.$$

Hence, we may write

$$\operatorname{Re}\varphi(s_k) = c_0\sigma_k + \operatorname{Re}(c_1) + \sum_{n=2}^{N} n^{-\sigma_k}(\operatorname{Re}(c_n)a_n + \operatorname{Im}(c_n)b_n) + \sum_{n=2}^{N} n^{-\sigma_k}F_n(t_k),$$

where each $F_n(t_k) \to 0$ as $k \to \infty$. Since $\operatorname{Re} s_k = \sigma_k$ also goes to 0, we may deduce

$$\operatorname{Re}(c_1) + \sum_{n=2}^{N}(\operatorname{Re}(c_n)a_n + \operatorname{Im}(c_n)b_n) = 0,$$

so that we have

$$\operatorname{Re}\varphi(s) = c_0\sigma + \sum_{n=2}^{N}(n^{-\sigma} - 1)(\operatorname{Re}(c_n)a_n + \operatorname{Im}(c_n)b_n) + \sum_{n=2}^{N} n^{-\sigma}F_n(t).$$

We will now choose another sequence $\{s'_k = \sigma'_k + it_k\}_{k\geq 1}$ where $\operatorname{Re}(s'_k) \to 0$ in order to obtain a contradiction with (7). More precisely, let $\{\sigma'_k\}_{k\geq 1}$ be any sequence of positive real numbers tending to 0 such that, for any $n = 2, \ldots, N$ and every $k \geq 1$, we have $n^{-\sigma'_k}|F_n(t_k)| \leq \sigma'_k$. Then we obtain

$$\begin{aligned}\operatorname{Re}\varphi(s'_k) &= c_0\sigma'_k + \sum_{n=2}^{N}(n^{-\sigma'_k} - 1)(\operatorname{Re}(c_n)a_n + \operatorname{Im}(c_n)b_n) + \sum_{n=2}^{N} n^{-\sigma'_k}F_n(t_k)\\ &= \mathbb{O}(\sigma'_k)\\ &= \mathbb{O}(\operatorname{Re}(s'_k)),\end{aligned}$$

and this contradicts (7). The assumption that φ has unrestricted range must be wrong. □

Remark. An inspection of the proof reveals that the statement of Theorem 1 remains true if we assume that $\varphi(s) = c_0 s + c_1 + \sum_{n=2}^{\infty} c_n n^{-s} \in \mathcal{G}$ with $c_0 \geq 1$, $\sum_{n=1}^{\infty}|c_n| < +\infty$ and that the complex dimension of φ is finite. The latter assumption is needed to use (7).

3. Preliminaries

As explained in the introduction, our main tool for proving or disproving compactness is a result from [Queffélec and Seip 2015a]. We formulate it in a more general context than for polynomials since it will be used under this form in Section 8. Recall that a *Carleson square* in $\mathbb{C}_0$ is a closed square in $\overline{\mathbb{C}_0}$ with one of its sides lying on the vertical line $i\mathbb{R}$; the side length of Q is denoted by $\ell(Q)$. A nonnegative Borel measure μ on $\overline{\mathbb{C}_0}$ is called a *vanishing Carleson measure* if

$$\limsup_{\ell(Q)\to 0} \frac{\mu(Q)}{\ell(Q)} = 0.$$

Lemma 5. *Suppose that $\varphi(s) = \sum_{n\geq 1} c_n n^{-s} \in \mathscr{G}$ and that $\varphi(\mathbb{C}_0)$ is bounded. The corresponding composition operator $\mathscr{C}_\varphi$ is compact on $\mathscr{H}^2$ if and only if the measure*

$$\mu_\varphi(E) := \boldsymbol{m}_\infty(\{z \in \mathbb{T}^\infty : \Phi(z) \in E\}), \qquad E \subseteq \mathbb{C}_0$$

is vanishing Carleson in $\mathbb{C}_0$, where Φ denotes a Bohr lift of φ.

Proof. This is Corollary 4.1 in [Queffélec and Seip 2015a]. □

Hence we consider squares

$$Q = Q(\tau, \varepsilon) = [0, \varepsilon] \times [\tau - \varepsilon/2, \tau + \varepsilon/2],$$

and want to investigate whether $\mu_\varphi(Q) = o(\varepsilon)$ uniformly in $\tau \in \mathbb{R}$. Our next lemma points out that this depends only on the local behavior of Φ.

Lemma 6. *Let φ be a Dirichlet polynomial* (2) *with $c_0 = 0$ mapping $\mathbb{C}_0$ into $\mathbb{C}_{1/2}$ and let Φ be a minimal Bohr lift of φ. If for every $w \in \mathbb{T}^d$ with* $\operatorname{Re}\Phi(w) = 0$ *there exists a neighborhood $\mathscr{U}_w \ni w$ in $\mathbb{T}^d$, constants $C_w > 0$ and $\kappa_w > 1$ such that for every $\tau \in \mathbb{R}$ and every $\varepsilon > 0$ we have*

$$\boldsymbol{m}_d(\{z \in \mathscr{U}_w : \Phi(z) \in Q(\tau, \varepsilon)\}) \leq C_w \varepsilon^{\kappa_w}, \tag{8}$$

then $\mathscr{C}_\varphi$ is compact.

Proof. Since φ is a Dirichlet polynomial, it has finite complex dimension d.

We first observe that (8) is always satisfied for those $w \in \mathbb{T}^d$ with $\operatorname{Re}\Phi(w) > 0$. Indeed, by continuity of Φ, we may always find a neighborhood $\mathscr{U}_w \ni w$ and $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$ and all $\tau \in \mathbb{R}$, $\{z \in \mathscr{U}_w : \Phi(z) \in Q(\tau, \varepsilon)\}$ is empty. We may then take $\kappa_w > 1$ be arbitrary and choose C_w with $C_w \varepsilon_0^{\kappa_w} \geq 1$.

We will then use a compactness argument and Lemma 5. Indeed, there exists a finite number of points $w_1, \ldots, w_N$ such that $\mathbb{T}^d$ is covered by $\mathscr{U}_{w_1}, \ldots, \mathscr{U}_{w_N}$. Now, we may take $C = C_{w_1} + \cdots + C_{w_N}$ and $\kappa = \min(\kappa_{w_1}, \ldots, \kappa_{w_N})$. Hence, for all $\tau \in \mathbb{R}$ and all $\varepsilon > 0$,

$$\boldsymbol{m}_d(\{z \in \mathbb{T}^d : \Phi(z) \in Q(\tau, \varepsilon)\}) \leq C\varepsilon^\kappa,$$

which achieves the proof of the compactness of $\mathscr{C}_\varphi$ on $\mathscr{H}^2$. □

Hence, we will require more information about the Taylor coefficients of Φ at a boundary point. Assume that $\Phi(w) = 0$ where $w = \mathbf{1}$. In this case, we will rewrite

$$\Phi(z) = \sum_{n\in\Lambda} \widetilde{c_n} \prod_{j=1}^{d} (1 - z_j)^{\alpha_j} = \sum_{\alpha\in\mathbb{N}^d} c_\alpha (1 - z)^\alpha, \tag{9}$$

where we have adopted the convention $c_\alpha = \widetilde{c_n}$, which is not generally equal to c_n. We shall need a kind of Julia–Carathéodory theorem for Φ of the form (9).

Lemma 7. *Let $\Phi : \mathbb{D}^d \to \mathbb{C}_0$ be of the form* (9) *and let $|\alpha| = 1$. Then $c_\alpha \geq 0$. Moreover, there exists at least one multi-index α with $|\alpha| = 1$ and $c_\alpha > 0$, unless $\Phi \equiv 0$.*

Proof. We may assume that $\alpha = (1, 0, \ldots, 0)$. Consider the one-variable polynomial

$$\psi(w) = \Phi(w, 1, \ldots, 1).$$

Clearly, ψ maps $\mathbb{D}$ to $\mathbb{C}_0$, and $\psi(1) = 0$. We write

$$\psi(w) = a(1 - w) + b(1 - w)^2 + \mathbb{O}((1 - w)^3).$$

We set $w = e^{i\theta}$ and obtain

$$\psi(e^{i\theta}) = a\Big(\frac{\theta^2}{2} - i\theta\Big) - b\theta^2 + \mathbb{O}(\theta^3).$$

In particular,

$$\mathrm{Re}(\psi(e^{i\theta})) = \theta\,\mathrm{Im}(a) + \theta^2\Big(\frac{\mathrm{Re}(a)}{2} - \mathrm{Re}(b)\Big) + \mathbb{O}(\theta^3).$$

Since this should be nonnegative, clearly $\mathrm{Im}(a) = 0$. We now set $w = 1 - \delta$ for $0 < \delta < 1$ and consider $\psi(\delta) = a\delta + \mathbb{O}(\delta^2)$. Since the real part of this also should be nonnegative as $\delta \to 0^+$ we must have $a \geq 0$. Hence $c_\alpha \geq 0$ when $|\alpha| = 1$.

Now, consider the mapping

$$\alpha \mapsto n(\alpha) = \prod_{j=1}^{d} p_j^{\alpha_j}.$$

It defines a total order on $\mathbb{N}^d$ by setting $\alpha \leq \beta$ if and only if $n(\alpha) \leq n(\beta)$. Assume that $\Phi \not\equiv 0$ and that $c_\alpha = 0$ whenever $|\alpha| = 1$. Consider

$$\beta = \inf\{\alpha : c_\alpha \neq 0\},$$

which exists since $\Phi \not\equiv 0$. There is $\theta \in (-\pi, \pi]$ so that $c_\beta = |c_\beta| e^{i\theta}$. Fix $\theta_j \in \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ and define

$$z_j = 1 - p_j^{-\sigma} e^{i\theta_j},$$

where $\sigma > 0$. For large enough σ, clearly $z = (z_1, \dots, z_d) \in \mathbb{D}^d$. Moreover,

$$\Phi(z_1, \dots, z_d) = |c_\beta| e^{i\theta}[n(\beta)]^{-\sigma} e^{i(\beta_1\theta_1+\cdots+\beta_d\theta_d)} + o([n(\beta)]^{-\sigma}),$$

as $\sigma \to \infty$. This implies that

$$\mathrm{Re}(\Phi(z_1, \dots, z_d)) = |c_\beta|[n(\beta)]^{-\sigma} \cos(\theta + \beta_1\theta_1 + \cdots \beta_d\theta_d) + o([n(\beta)]^{-\sigma}).$$

Since $|\beta| \geq 2$, we can choose $\theta_j \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ such that $\cos(\theta+\beta_1\theta_1+\cdots+\beta_d\theta_d) < 0$. This contradicts the mapping properties of Φ, and hence the assumption that $c_\alpha = 0$ whenever $|\alpha| = 1$ is wrong. □

We will also need two lemmas from differential geometry. The first one is the parametrized Morse lemma (see for instance [Bruce and Giblin 1992, § 4.44]).

Lemma (parametrized Morse lemma). *Let $\mathcal{U} \subset \mathbb{R}^J \times \mathbb{R}^{d-J}$ be a neighborhood of $\mathbf{0} \in \mathbb{R}^d$ and let $F : \mathcal{U} \to \mathbb{R},\ (u, v) \mapsto F(u, v)$ be a smooth function. Assume that $F(\mathbf{0}) = 0$, that $\partial F/\partial u_i(\mathbf{0}) = 0$ for all $i = 1, \dots, J$ and that the matrix*

$$\left(\frac{\partial^2 F}{\partial u_i \partial u_j}(\mathbf{0})\right)_{1\leq i,\, j\leq J}$$

is positive definite. Then there exist a neighborhood $\mathcal{V} \ni \mathbf{0}$ with $\mathcal{V} \subset \mathcal{U}$, a smooth diffeomorphism $\Gamma : \mathcal{V} \to \mathbb{R}^d,\ (u, v) \mapsto (\gamma(u, v), v)$ with $\Gamma(\mathbf{0}) = \mathbf{0}$ and a smooth map $h : \mathbb{R}^{d-J} \to \mathbb{R}$ such that, for any $(u, v) \in \mathcal{V}$,

$$F(u, v) = \sum_{j=1}^{J} \gamma_j(u, v)^2 + h(v).$$

The second lemma reads as follows:

Lemma 8. *Let $p \geq 1$ be an integer, and let $f : I \to \mathbb{R}$ be a smooth function where I is an open interval containing 0 and $f(x) \sim_0 x^p$. Then there exist $C > 0$ and an open interval $I' \ni 0$ inside I such that, for any $\tau \in \mathbb{R}$ and any $\delta > 0$, the set $\{x \in I' : |f(x) - \tau| < \delta\}$ has Lebesgue measure less than $C\delta^{1/p}$.*

Proof. Assume first that $f(x) = x^p$. If $|\tau| \leq 2\delta$, then the result is clear. Otherwise, if $\tau \geq 2\delta$, then x has to live in $[(\tau-\delta)^{1/p}, (\tau+\delta)^{1/p}]$ and the length of this interval may be easily estimated using the mean value theorem.

The general case reduces to this one. For small values of x, set $y = [f(x)]^{1/p}$ if p is odd or $y = [f(x)]^{1/p}$ for $x > 0$, $y = -[f(x)]^{1/p}$ for $x < 0$ if p is even. In both cases, y is differentiable at 0 and $dy/dx > 0$. Hence, $x = \gamma(y)$ where γ is a smooth diffeomorphism. Now, for some small open interval $I' \ni 0$, we have

$$\{x \in I' : |f(x) - \tau| < \delta\} = \{x \in I' : |(\gamma^{-1}(x))^p - \tau| < \delta\}.$$

Since γ is a diffeomorphism, the latter set has Lebesgue measure less than $C\delta^{1/p}$. □

4. Proof of Theorem 2

We intend to apply Lemma 6. Hence, let $w \in \mathbb{T}^d$ with $\operatorname{Re}\Phi(w) = 0$. By the rotational invariance of $\boldsymbol{m}_d$, we may always assume that $w = \mathbf{1}$. Moreover, since the conditions in Lemma 6 are invariant by vertical translations, we may also assume that $\Phi(w) = 0$. In this case we have

$$\Phi(z_1, z_2, \dots, z_d) = \sum_{j=1}^{d} \Phi_j(z_j) = \sum_{j=1}^{d} \sum_{k} a_k^{(j)} (1 - z_j)^k.$$

Since Φ is a minimal Bohr lift of φ, inspecting the proof of Lemma 7, we may conclude that in this case $a_1^{(j)} > 0$ for every $j = 1, 2, \dots, d$. This means we have

$$\operatorname{Re}\Phi(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_d}) = \sum_{j=1}^{d} b_j \theta_j^{k_j} + o\big(\theta_j^{k_j}\big),$$

where the coefficients $b_j \neq 0$ are real numbers and the exponents $k_j \geq 2$ are integers. The fact that this quantity is supposed to be nonnegative implies that $b_j > 0$ and that k_j is even, by similar considerations as those in the proof of Lemma 7. Moreover

$$\operatorname{Im}\Phi(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_d}) = -\sum_{j=1}^{d} a_1^{(j)} \theta_j + o(\theta_j).$$

Proof of the first part of Theorem 2. Let $\tau \in \mathbb{R}$ and $\varepsilon > 0$ be arbitrary. The preceding discussion means there is some neighborhood $\mathcal{U} \ni (1, 1, \dots, 1)$ in $\mathbb{T}^d$ so that

$$\frac{1}{2} \sum_{j=1}^{d} b_j \theta_j^{k_j} \leq \operatorname{Re}\Phi(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_d}) \leq 2 \sum_{j=1}^{d} b_j \theta_j^{k_j},$$

when $e^{i\theta} \in \mathcal{U}$. Hence if $\Phi(e^{i\theta}) \in Q(\tau, \varepsilon)$ and $e^{i\theta} \in \mathcal{U}$, we conclude from the real part that $|\theta_j| \ll \varepsilon^{1/k_j}$, for $j = 1, 2, \dots, d$. Now, fixing θ_j for $j = 2, \dots, d$, we conclude from the imaginary part and Lemma 8 that θ_1 can live in an interval of size at most $C\varepsilon$. Hence we have

$$\boldsymbol{m}_d(\{z \in \mathcal{U}_w : \Phi(z) \in Q(\tau, \varepsilon)\}) \ll_w \varepsilon^{1+1/k_2+\cdots+1/k_d}.$$

In fact, we may choose

$$\kappa_w = 1 + \sum_{j=1}^{d} \frac{1}{k_j} - \min_{1 \leq j \leq d} \frac{1}{k_j},$$

and conclude by Lemma 6, since $d \geq 2$. $\square$

Proof of the second part of Theorem 2. In this case $d = 1$, and the polynomial $\Phi(z)$ is of only one variable. We again consider some neighborhood $\mathcal{U} \ni 1$ in $\mathbb{T}$, so that when $e^{i\theta} \in \mathcal{U}$ we have

$$0 \le \operatorname{Re} \Phi(e^{i\theta}) \le 2b\theta^k \quad \text{and} \quad |\operatorname{Im} \Phi(e^{i\theta})| \le 2a|\theta|,$$

where $a = a_1,\ b = b_1$ and $k \ge 2$ is even. Now, we choose $\tau = 0$ and observe that $\varphi(e^{i\theta})$ belongs to $Q(\tau, \varepsilon)$ provided $|\theta| \ll \varepsilon$. Hence

$$\boldsymbol{m}_1(\{z \in \mathbb{T} : \Phi(z) \in Q(\tau, \varepsilon)\}) \ge \boldsymbol{m}_1(\{z \in \mathcal{U}_w : \Phi(z) \in Q(\tau, \varepsilon)\}) \gg \varepsilon,$$

and $\mathcal{C}_\varphi$ cannot be compact by Lemma 5. □

Remark. Inspecting the proof of Theorem 2, we see that we may replace the polynomials P_j, by corresponding power series

$$P_j(q_j^{-s}) = \sum_{k=0}^{\infty} c_k^{(j)} q_j^{-ks},$$

provided $\sum_{k=0}^{\infty} |c_k^{(j)}| < \infty$. However, we still require the complex dimension d to be finite.

5. Proof of Theorem 3

We begin by observing that the penultimate point of Theorem 3 follows from the second part of Theorem 2. The final part of Theorem 3 is contained in the following result:

Lemma 9. *There are polynomials $\varphi \in \mathcal{G}$ of any complex dimension and of any degree ≥ 3 for which the corresponding composition operator $\mathcal{C}_\varphi$ is noncompact.*

Proof. Let $P(z) = P(z_1, z_2, \dots, z_d)$ be any polynomial in d variables and define

$$\Phi(z) = (1 - z_1) + \delta(1 - z_1)^2 P(z),$$

for some $\delta > 0$ to be decided later. We compute

$$\operatorname{Re} \Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) = (1-\cos\theta_1)(1-2\delta(\cos(\theta_1) \operatorname{Re} P(e^{i\theta}) - \sin(\theta_1) \operatorname{Im} P(e^{i\theta}))).$$

Pick δ small enough so that we have

$$\frac{1 - \cos\theta_1}{2} \le \operatorname{Re} \Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) \le 2(1 - \cos\theta_1).$$

The first inequality tells us that Φ is a minimal Bohr lift of

$$\varphi(s) = (1 - p_1{}^{-s}) + \delta(1 - p_1{}^{-s})^2 P(p_1^{-s}, \dots, p_d^{-s}),$$

with $\varphi \in \mathcal{G}$ having unrestricted range. Using the second inequality and a Taylor expansion of $\operatorname{Im}\Phi$, we also get that near $\mathbf{1}$,

$$\operatorname{Re}\Phi(e^{i\theta_1},\dots,e^{i\theta_d}) = \mathbb{O}(\theta_1^2) \quad \text{and} \quad \operatorname{Im}\Phi(e^{i\theta_1},\dots,e^{i\theta_d}) = \mathbb{O}(\theta_1).$$

Similar considerations as in the proof of the second part of Theorem 2 allow us to conclude that $\mathscr{C}_\varphi$ is not compact. □

Remark. The key point of Lemma 9 is that even if Φ involves d variables, its local behavior near $\mathbf{1}$ depends too heavily on z_1 to ensure compactness.

Having now concluded the negative parts of Theorem 3, we turn to the positive parts. Let us fix a polynomial $\varphi \in \mathcal{G}$ and let Φ denote a minimal Bohr lift of φ. We can simplify how to write Φ around a point $w \in \mathbb{T}^d$ such that $\operatorname{Re}\Phi(w) = 0$. Without loss of generality, we may again assume that $w = \mathbf{1}$ and that $\Phi(w) = 0$. Then we may write

$$\Phi(z)=\sum_{j=1}^{d}a_j(1-z_j)+\sum_{j=1}^{d}b_j(1-z_j)^2+\sum_{1\le j<k\le d}c_{j,k}(1-z_j)(1-z_k)+o\biggl(\sum_{1\le j\le d}|1-z_j|^2\biggr).$$

We let $z_j = e^{i\theta_j}$ and since $a_j \ge 0$ by Lemma 7 we get

$$\operatorname{Re}(\Phi(z)) = \sum_{j=1}^{d}\Bigl(\frac{a_j}{2}-\operatorname{Re}(b_j)\Bigr)\theta_j^2 - \sum_{1\le j<k\le d}\operatorname{Re}(c_{j,k})\theta_j\theta_k + o\biggl(\sum_{1\le j\le d}\theta_j^2\biggr).$$

The quadratic form appearing above is brought to standard form by a linear change of variables,

$$\operatorname{Re}(\Phi(z)) = \sum_{j=1}^{d}(\ell_j(\theta))^2 + o\biggl(\sum_{1\le j\le d}\theta_j^2\biggr).$$

Next, we write

$$\operatorname{Im}(\Phi(z)) = -\sum_{j=1}^{d}a_j\theta_j + o\biggl(\sum_{j=1}^{d}|\theta_j|\biggr) = -\ell_{d+1}(\theta) + o\biggl(\sum_{j=1}^{d}|\theta_j|\biggr),$$

and by Lemma 7 we know that $\ell_{d+1} \not\equiv 0$, since at least one $a_j > 0$. The last step to finish the proof of Theorem 3 is the following result:

Lemma 10. *Let $\Phi : \mathbb{D}^d \to \mathbb{C}_0$ be an optimal Bohr lift of $\varphi \in \mathcal{G}$, where φ has unrestricted range and complex dimension $d \ge 2$. Suppose that $w \in \mathbb{T}^d$ is such that* $\operatorname{Re}\Phi(w) = 0$. *Then there exist a neighborhood $\mathcal{U}_w \ni w$ in $\mathbb{T}^d$, $\kappa = \kappa_w > 1$ and $C = C_w > 0$ such that, for any $\tau \in \mathbb{R}$ and for every $\varepsilon > 0$,*

$$\boldsymbol{m}_d(\{z \in \mathcal{U}_w : \Phi(z) \in Q(\tau,\varepsilon)\}) \le C\varepsilon^{\kappa}.$$

In the following cases:

- *$J(\Phi, w) \geq 1$ and ℓ_{d+1} is independent from $(\ell_1, \ldots, \ell_J)$. We may choose $\kappa = 1 + J(\Phi, w)/2$.*
- *$J(\Phi, w) \geq 2$ and ℓ_{d+1} belongs to $\operatorname{span}(\ell_1, \ldots, \ell_J)$. We may choose $\kappa = (1 + J(\Phi, w))/2$.*
- *$J(\Phi, w) = 1$, ℓ_{d+1} is a multiple of ℓ_1 and Φ has degree 2. We may choose $\kappa = \frac{9}{8}$.*
- *$J(\Phi, w) = 0$ and Φ has degree 2. We may choose $\kappa = (d+3)/4$.*

Before we prove the different cases of this lemma, let us make some comments. Firstly, it is clear that Lemma 10 and Lemma 6 imply Theorem 3 when the degree of φ is at least 2. When the degree of φ is equal to 1, then

$$\Phi(z) = \sum_{j=1}^{d} a_j(1 - z_j)$$

so that each a_j is positive. This implies that $J(\Phi, w) = d$ so that we may again apply Lemma 10 and Lemma 6.

It is also important to notice that Φ cannot be an arbitrary polynomial mapping of $\mathbb{D}^d$ into $\mathbb{C}_0$. It is an optimal Bohr lift of some $\varphi \in \mathscr{G}$ with complex dimension d. In particular, we shall use that $\partial\Phi/\partial z_j \not\equiv 0$ for every $1 \leq j \leq d$. Moreover, the polynomial $\Phi(z) = \lambda(1 - z_1 z_2)$ is not an optimal Bohr lift. Otherwise, it would arise from $\varphi(s) = \lambda(1 - q_1^{-s} q_2^{-s})$, but the optimal Bohr lift of φ is $\lambda(1 - z)$.

We are now ready for the proof of Lemma 10. By similar considerations as before, we may again assume that $w = \mathbf{1}$ and that $\Phi(w) = 0$. We write J for $J(\Phi, w)$.

The case $J = 0$. This implies that

$$\frac{a_j}{2} - \operatorname{Re}(b_j) = \operatorname{Re}(c_{j,k}) = 0$$

for $j, k = 1, \ldots, d$. We set $z_j = e^{i\theta_j}$ and compute

$$\begin{aligned}
\operatorname{Re}(a_j(1 - z_j)) &= a_j(1 - \cos\theta_j) \\
\operatorname{Re}(b_j(1 - z_j)^2) &= -a_j \cos\theta_j(1 - \cos\theta_j) + 2\operatorname{Im}(b_j)\sin\theta_j(1 - \cos\theta_j) \\
\operatorname{Re}(c_{j,k}(1 - z_j)(1 - z_k)) &= \operatorname{Im}(c_{j,k})(\sin\theta_j(1 - \cos\theta_k) + \sin\theta_k(1 - \cos\theta_j))
\end{aligned}$$

This means that

$$\operatorname{Re}(\Phi(z)) = \sum_{j=1}^{d} \operatorname{Im}(b_j)\theta_j^3 + \sum_{1 \leq j < k \leq d} \frac{\operatorname{Im}(c_{j,k})}{2}(\theta_j\theta_k^2 + \theta_k\theta_j^2) + o\left(\sum_{j=1}^{d} |\theta_j|^3\right).$$

However, the nonnegativity of $\operatorname{Re}\Phi$ then implies that $\operatorname{Im}(b_j) = \operatorname{Im}(c_{j,k}) = 0$. Hence we in total have $b_j = a_j/2$ and $c_{j,k} = 0$, which means

$$\Phi(z) = \sum_{j=1}^{d} \Big(a_j(1 - z_j) + \frac{a_j}{2}(1 - z_j)^2 \Big).$$

In fact, this means that $a_j > 0$ for every j, by the assumption that the complex dimension is d and Lemma 7. We may now use (the proof of) Theorem 2 to conclude that there exists a neighborhood $\mathcal{U}_w \ni w$ such that

$$\boldsymbol{m}_d(\{z \in \mathcal{U}_w : \Phi(z) \in Q(\tau, \varepsilon)\}) \ll \varepsilon \times \varepsilon^{\frac{d-1}{4}} = \varepsilon^{\frac{d+3}{4}},$$

since we now have

$$\operatorname{Re}\Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) = \frac{1}{4}\sum_{j=1}^{d} a_j \theta_j^4 + o(\theta_j^4),$$

and we are done with this case. □

The case $J \geq 1$ and independence. After a linear change of variables, we may write $\operatorname{Re}\phi$ and $\operatorname{Im}\phi$ as

$$\operatorname{Re}\phi(\theta_1, \dots, \theta_d) = u_1^2 + \dots + u_J^2 + o\Big(\sum_{j=1}^{d} u_j^2\Big),$$

$$\operatorname{Im}\phi(\theta_1, \dots, \theta_d) = u_d + o\Big(\sum_{j=1}^{d} |u_j|\Big).$$

Since a linear change of variables does not change the value of the volume up to constants, we may assume that ϕ depends on (u, v) with $u = (u_1, u_2, \dots, u_J)$ and $v = (u_{J+1}, \dots, u_d)$. Applying the parametrized Morse lemma to $\operatorname{Re}\phi$, we may write

$$\operatorname{Re}\phi(u, v) = \gamma_1(u, v)^2 + \dots + \gamma_J(u, v)^2 + h(v).$$

We also apply the change of variables $(u,v) \mapsto \Gamma(u,v)$ to $\operatorname{Im}\phi$ and since $\Gamma_d(u,v) = u_d$, we find

$$\operatorname{Im}\phi(u, v) = u_d + g(\Gamma(u, v)),$$

where g is a smooth function defined on $\mathcal{V}$ such that $\partial g/\partial u_d(\mathbf{0}) = 0$.

Now, we know that $\operatorname{Re}\phi(u, v) \geq 0$ for every $(u, v) \in \mathbb{R}^d$. Since Γ is a diffeomorphism, v can take any value in some neighborhood of zero in $\mathbb{R}^{d-J}$ even if we require that

$$\gamma_1(u, v) = \gamma_2(u, v) = \dots = \gamma_J(u, v) = 0,$$

and hence $h(v) \geq 0$.

This implies that we may find some neighborhood $\mathcal{W} \ni \mathbf{0}$ in $\mathcal{V}$ such that, for every $\tau \in \mathbb{R}$ and every $\varepsilon > 0$,

$$(u, v) \in \mathcal{W} \quad \text{and} \quad \phi(u, v) \in Q(\tau, \varepsilon) \qquad \Longrightarrow \qquad |\gamma_j(u, v)| \leq \varepsilon^{1/2}.$$

Now, for if we fix $\gamma_1(u, v), \dots, \gamma_{d-1}(u, v)$, it follows from Lemma 8 with $p = 1$ that $\gamma_d(u, v) = u_d$ has to belong to some interval of size $C\varepsilon$, provided that (u, v) is sufficiently close to $\mathbf{0}$. This means that there exists a neighborhood $\mathcal{O} \subset \mathcal{W}$ of $\mathbf{0}$ such that

$$\{(u, v) \in \mathcal{O} : \phi(u, v) \in Q(\tau, \varepsilon)\} \subset \{(u, v) \in \mathcal{O} : \Gamma(u, v) \in R(\tau, \varepsilon)\},$$

where the volume of $R(\tau, \varepsilon)$ is less than $C\varepsilon^{1+(J/2)}$. Since Γ is a diffeomorphism, we are done. □

The case $J \geq 2$ and dependence. With a similar linear change of variables as in the previous case, we may write

$$\operatorname{Re}\phi(u_1, \dots, u_d) = u_1^2 + \cdots + u_J^2 + o\biggl(\sum_{j=1}^{d} u_j^2\biggr),$$

$$\operatorname{Im}\phi(u_1, \dots, u_d) = \sum_{j=1}^{J} \alpha_j u_j + o\biggl(\sum_{j=1}^{d} |u_j|\biggr).$$

We use again the parametrized Morse lemma with $\operatorname{Re}\phi$, and it is again easy to show that $\gamma_j(u, v) = u_j + o\bigl(\sum_{j=1}^{d} |u_j|\bigr)$ so that

$$\operatorname{Im}\phi(u, v) = \sum_{j=1}^{J} \alpha_j \gamma_j(u, v) + g(\Gamma(u, v))$$

with $\partial g/\partial u_j(\mathbf{0}) = 0$ for $j = 1, \dots, d$.

We argue as in the previous case. For every $j = 2, \dots, J$, for any $\tau \in \mathbb{R}$ and every $\varepsilon > 0$,

$$(u, v) \in \mathcal{W} \subset \mathcal{V} \quad \text{and} \quad \phi(u, v) \in Q(\tau, \varepsilon) \qquad \Longrightarrow \qquad |\gamma_j(u, v)| \leq \varepsilon^{1/2}.$$

Now, for a fixed value of $\gamma_2(u, v), \dots, \gamma_d(u, v)$, it is again clear that $\gamma_1(u, v)$ has to belong to some interval of size $C\varepsilon$, provided (u, v) is sufficiently close to $\mathbf{0}$. This means that there exists a neighborhood $\mathcal{O} \ni \mathbf{0}$ in $\mathcal{W}$ such that

$$\{(u, v) \in \mathcal{O} : \phi(u, v) \in Q(\tau, \varepsilon)\} \subset \{(u, v) \in \mathcal{O} : \Gamma(u, v) \in R(\tau, \varepsilon)\},$$

where the volume of $R(\tau, \varepsilon)$ is less than $C\varepsilon^{1+(J-1/2)}$. We conclude as in the previous step. □

The case $J = 1$ and dependence, $d = 2$. This is the most difficult case. At first, we do not assume that $d = 2$ but we always assume that the degree of φ is equal to 2. We know that there is constant $\lambda \in \mathbb{R}^*$ so that $\ell_1(\theta) = \lambda \ell_{d+1}(\theta)$, which means

$$\sqrt{\frac{a_j}{2} - \operatorname{Re}(b_j)} = \lambda a_j, \qquad 1 \le j \le d$$

and that $\lambda > 0$ by the computations in the beginning of this section. We normalize $\Phi(z)$ as $\lambda^{-2}\Phi(z)$, so that we may assume that $\lambda = 1$. Hence

$$\ell_1(\theta) = \sum_{j=1}^{d} a_j \theta_j,$$

and this immediately implies that

$$\operatorname{Re}(b_j) = \frac{a_j}{2} - a_j^2 \quad \text{and} \quad \operatorname{Re}(c_{j,k}) = -2a_j a_k, \quad 1 \le j,\ k \le d. \tag{10}$$

Suppose that $a_1 = 0$. Then $\operatorname{Re}(b_1) = 0$ and $\operatorname{Re}(c_{1,k}) = 0$ for $2 \le k \le d$. We compute

$$\operatorname{Re}(\Phi(e^{ix}, 1, \ldots, 1)) = -2\operatorname{Im}(b_1)\sin x(1 - \cos x) \ge 0,$$

which means that $\operatorname{Im}(b_1) = 0$, so that $b_1 = 0$. Next we compute

$$\begin{aligned}
\Phi(e^{ix}, e^{iy}, 1\ldots, 1) = a_2(1 - \cos y) + \Big(\frac{a_2}{2} - a_2^2\Big)&(1 - 2\cos y + \cos 2y) \\
&- \operatorname{Im}(c_{1,2})(-\sin x - \sin y + \sin(x + y)) \\
= (1 - \cos y)(a_2(1 - \cos y) &+ 2a_2^2 \cos y + \operatorname{Im}(c_{1,2}) \sin x) \\
&+ \operatorname{Im}(c_{1,2}) \sin y(1 - \cos x).
\end{aligned}$$

Taking $y = \pm\delta$ for small enough δ, we obtain that $\operatorname{Im}(c_{1,2}) = 0$. There is nothing special about z_2, and hence we conclude that $\operatorname{Im}(c_{1,k}) = 0$, for $2 \le k \le d$. In particular, $c_{1,k} = 0$ for the same values of k. But this is impossible, since the variable z_1 no longer appears in our polynomial. Hence the assumption that $a_1 = 0$ must be wrong.

Arguing in the same way, we have that $a_j > 0$ for $1 \le j \le d$. Moreover, after renaming the variables, we may suppose $a_1 \ge a_2 \ge \cdots a_d > 0$. Finally,

$$0 \le \operatorname{Re}(\Phi(-1, 1, \ldots, 1)) = 2a_1 + 4\Big(\frac{a_1}{2} - a_1^2\Big) \qquad \Longrightarrow \qquad a_1 \le 1,$$

so without loss of generality, we may assume that

$$1 \ge a_1 \ge a_2 \ge \cdots \ge a_d > 0.$$

From now on, we assume that $d = 2$ and that $1 \ge a_1 \ge a_2 > 0$. We need the following lemma.

Lemma 11. *We have* $a_2 \leq 1 - a_1$.

Proof. We compute

$$\Phi(-1,-1) = -4a_1^2 - 4a_2^2 - 8a_1a_2 + 4a_1 + 4a_2 = 4(a_1+a_2)(1-a_1-a_2).$$

Since this has to be nonnegative, we get the result. □

Remark. Lemma 11 immediately implies that $a_1 \in (0,1)$ and $a_2 \in \left(0, \frac{1}{2}\right]$ by the assumptions that $0 < a_2 \leq a_1 \leq 1$.

Let us apply the change of variables $\theta_1 = a_2u + a_2v,\ \theta_2 = a_1u - a_1v$ to ϕ:

$$\operatorname{Re}\phi(u,v) = -4a_1^2a_2^2u^2 + o(u^2+v^2), \tag{11}$$

$$\operatorname{Im}\phi(u,v) = 2a_1a_2u + o(|u|+|v|). \tag{12}$$

As before, we intend to apply the parametrized Morse lemma to $\operatorname{Re}\phi$. Setting $\Psi = \Gamma^{-1}$, we get that, around $\mathbf{0}$,

$$\operatorname{Re}\phi\circ\Psi(u,v) = u^2 + h(v) \quad \text{and} \quad \operatorname{Im}\phi\circ\Psi(u,v) = u + g(u,v),$$

with h and g smooth functions which have no terms of order 1 at $\mathbf{0}$.

Assume first that $h \not\equiv 0$. Let $p \geq 2$ be such that $h(v) \sim_0 \alpha_p v^p$ with $\alpha_p \neq 0$. Because $\phi\circ\Psi$ maps $\mathbb{R}^2$ into $\overline{\mathbb{C}_0}$, we must have that $\alpha_p > 0$ and that p is even. Now, if $\phi\circ\Psi(u,v) \in Q(\tau,\varepsilon)$ with (u,v) sufficiently close to $\mathbf{0}$, then $0 \leq h(v) \leq \varepsilon$ which implies by Lemma 8 that v belongs to some set of measure less than $C\varepsilon^{1/p}$. Moreover, for a fixed value of v, a look at the imaginary part and Lemma 8 yield that u has to belong to some interval of size $C\varepsilon$ and thus we are done with $\kappa = 1 + 1/p$.

Thus, we are lead to study what happens if $h \equiv 0$. The situation is easier if the Taylor expansion of $g(u,v)$ admits some term in v^p. In that case, we may write

$$\operatorname{Im}\phi\circ\Psi(u,v) = ug_1(u,v) + v^pg_2(v),$$

with smooth functions g_1 and g_2, such that $g_1(0,0)=1$ and $g_2(0)\neq 0$. If $\phi\circ\Psi(u,v)$ belongs to $Q(\tau,\varepsilon)$, we conclude from the real part that then $|u| \leq \varepsilon^{1/2}$, and from the imaginary part, we get that, near $\mathbf{0}$,

$$|v^pg_2(v) - \tau| \leq C\varepsilon^{1/2}.$$

By appealing again to Lemma 8, we conclude that v belongs to some set of Lebesgue measure less than $C\varepsilon^{1/2p}$. For a fixed value of v, we look once more at the imaginary part, and obtain that u must belongs to some interval of size $C\varepsilon$. Hence, we are done with $\kappa = 1 + 1/(2p)$.

Therefore, it remains to show that we will always fall into one of the previous cases and compute the value of p. We again recall that the polynomial

$$\Phi(z) = \lambda(1 - z_1z_2),$$

is a contradiction to the fact that Φ is a minimal Bohr lift of $\varphi \in \mathcal{G}$. More precisely, we are reduced to proving the following lemma.

Lemma 12. *Let* $0 < a_2 \leq a_1 \leq 1$ *and* $a_2 \leq 1 - a_1$. *Suppose that* $\Phi : \mathbb{D}^2 \to \mathbb{C}_0$ *is the polynomial*

$$\Phi(z) = a_1(1-z_1) + a_2(1-z_2) + b_1(1-z_1)^2 + b_2(1-z_2)^2 + c(1-z_1)(1-z_2), \tag{13}$$

where

$$\mathrm{Re}(b_1) = \tfrac{1}{2}a_1 - a_1^2, \quad \mathrm{Re}(b_2) = \tfrac{1}{2}a_2 - a_2^2 \quad \textit{and} \quad \mathrm{Re}(c) = -2a_1a_2.$$

Set $\theta_1 = a_2u + a_2v,\ \theta_2 = a_1u - a_1v$ *and*

$$\phi(u, v) = \Phi(e^{i\theta_1}, e^{i\theta_2}).$$

Then there do not exist smooth maps $\gamma : \mathbb{R}^2 \to \mathbb{R}$ *and* $h : \mathbb{R}^2 \to \mathbb{R}$ *so that*

$$\mathrm{Re}\,\phi(u, v) = \gamma(u, v)^2, \tag{14}$$

$$\mathrm{Im}\,\phi(u, v) = \gamma(u, v)h(u, v), \tag{15}$$

except if $\Phi(z) = \frac{1}{2}(1 - z_1z_2)$. *More precisely, if* $\Phi(z) \neq \frac{1}{2}(1 - z_1z_2)$, *for any smooth maps* $\gamma : \mathbb{R}^2 \to \mathbb{R}$ *and* $h : \mathbb{R} \to \mathbb{R}$, *then*

- *either the Taylor series of* $\mathrm{Re}\,\phi - \gamma^2$ *at* $\mathbf{0}$ *has a nonzero term of order* ≤ 5,
- *or the Taylor series of* $\mathrm{Im}\,\phi - \gamma \cdot h$ *at* $\mathbf{0}$ *has a nonzero term of order* ≤ 4.

The proof of this lemma is rather delicate and will be postponed to Section 6 in order to keep a clearer exposition of the proof of Lemma 10. However, using Lemma 12 we are able to finish this case. Indeed, if $\Gamma(u, v) = (\gamma(u, v), v)$ is the map given by the parametrized Morse lemma and if f_1, f_2 and f_3 are smooth functions such that

$$\mathrm{Re}\,\phi(u, v) = \gamma(u, v)^2 + f_1(v) \quad \text{and} \quad \mathrm{Im}\,\phi(u, v) = \gamma(u, v)f_2(u, v) + f_3(v),$$

then Lemma 12 implies that either $f_1(v) \sim_0 \alpha_p v^p$ with $p \leq 5$ or $f_3(v) \sim_0 \beta_p v^p$ with $p \leq 4$. By the considerations above we conclude $\kappa = \frac{9}{8}$ is possible. $\square$

The case $J = 1$ ***and dependence,*** $d \geq 3$. We are left to consider the case $J = 1$, $d \geq 3$ and ℓ_1 is a multiple of ℓ_{d+1}. We shall deduce this case from the case $d = 2$ using the following lemma:

Lemma 13. *Let* $\varepsilon > 0,\ \tau \in \mathbb{R}$ *and* $z_3, \ldots, z_d \in \mathbb{T}^{d-2}$. *Consider the set*

$$A_{z_3,\ldots,z_d}(\tau, \varepsilon) = \{(z_1, z_2) \in \mathbb{T}^2 : \Phi(z) \in Q(\tau, \varepsilon)\}.$$

Then, for every $w_3, \dots, w_d \in \mathbb{T}^{d-2}$, there exists a neighborhood $\mathcal{W} \ni (w_3, \dots, w_d)$ in $\mathbb{T}^{d-2}$ such that, for all $(z_3, \dots, z_d) \in \mathcal{W}$ we have

$$A_{z_3,\dots,z_d}(\tau, \varepsilon) \subset A_{w_3,\dots,w_d}(\tau, 2\varepsilon).$$

Proof. Assume that this is not the case. Then there exists a sequence $(z_1^{(k)}, \dots, z_d^{(k)})$ in $\mathbb{T}^d$ such that $z_j^{(k)} \to w_j$ for $3 \le j \le d$ and

$$(z_1^{(k)}, z_2^{(k)}) \in A_{z_3^{(k)},\dots,z_d^{(k)}}(\tau, \varepsilon) \setminus A_{w_3,\dots,w_d}(\tau, 2\varepsilon).$$

Extracting a subsequence if necessary, we may assume that $z_1^{(k)} \to w_1$ and $z_2^{(k)} \to w_2$ for some $(w_1, w_2) \in \mathbb{T}^2$. By continuity of Φ, this implies $\Phi(w) \in \overline{Q(\tau, \varepsilon)} \setminus Q(\tau, 2\varepsilon)$, which is a contradiction. □

We now set $\Psi_{1,2}(z_1, z_2) = \Phi(z_1, z_2, 1, \dots, 1)$. Since $J(\Phi, w) = 1$, we already know that $a_j > 0$ for all $j = 1, \dots, d$ and hence the variables z_1 and z_2 both appear in the polynomial $\Psi_{1,2}$. Provided $\Psi_{1,2}(z) \neq \lambda_{1,2}(1 - z_1 z_2)$ for some $\lambda_{1,2} \in \mathbb{R}^*$, we know from the case $d = 2$ that there exists a neighborhood $\mathcal{V} \ni (z_1, z_2)$ in $\mathbb{T}^2$ and $C > 0$ such that, for any $\tau \in \mathbb{R}$ and every $\varepsilon > 0$,

$$\boldsymbol{m}_2(\{(z_1, z_2) \in \mathcal{V} : \Phi(z_1, z_2, 1, \dots, 1) \in Q(\tau, 2\varepsilon)\}) \le C\varepsilon^{\kappa}$$

with $\kappa = \frac{9}{8}$. By Lemma 13, there exists a neighborhood $\mathcal{W} \ni \mathbf{1}$ in $\mathbb{T}^{d-2}$ such that, for any $(z_3, \dots, z_d) \in \mathcal{W}$,

$$\{(z_1,z_2) \in \mathcal{V} : \Phi(z_1,\dots,z_d) \in Q(\tau,\varepsilon)\} \subset \{(z_1,z_2) \in \mathcal{V} : \Phi(z_1,z_2,1,\dots,1) \in Q(\tau,2\varepsilon)\}.$$

This yields $\boldsymbol{m}_d(\{z \in \mathcal{V} \times \mathcal{W} : \Phi(z) \in Q(\tau, \varepsilon)\}) \le C\varepsilon^{\kappa}$.

So the result is proved except if, for every $j < k$, there exists some $\lambda_{j,k} > 0$ such that

$$\begin{aligned}(16)\quad \Phi(1,\dots,1,z_j,1,\dots,1,z_k,1,\dots,1) &= \lambda_{j,k}(1-z_j z_k) \\ &= \lambda_{j,k}((1-z_j)+(1-z_k)-(1-z_j)(1-z_k)).\end{aligned}$$

Comparing this with the expansion of Φ near $\mathbf{1}$, we get

$$a_j = a_k = \lambda_{j,k}, \qquad b_j = 0, \qquad c_{j,k} = -\lambda_{j,k}.$$

Using (10), we may conclude that $a_j = \frac{1}{2}$ and $c_{j,k} = -\frac{1}{2}$. In total this means that

$$\Phi(z_1, z_2, \dots, z_d) = \frac{1}{2}\sum_{j=1}^{d}(1 - z_j) - \frac{1}{2}\sum_{1 \le j < k \le d}(1 - z_j)(1 - z_k).$$

However,

$$\Phi(-1, -1, \dots, -1) = \frac{2d - 2d(d-1)}{2} = d(2-d) < 0,$$

since $d \geq 3$. Hence (16) is not possible for every $j < k$ and we are done. $\square$

6. Proof of Lemma 12

We intend to prove this result by contradiction. We require several tedious computations, which can be done either by hand or by a computer algebra system. We have used Xcas, and our file is available for download [Bayart and Brevig 2015]. In the proof below, we will skip certain computations such as computing Taylor coefficients, simplifying algebraical expressions and solving simple equations. The proof consists of three steps, and in each step we refer to the lines in that file where the computations are performed.

The idea of the argument is rather easy. We assume that we may factorize $\operatorname{Re}\phi(u, v)$ and $\operatorname{Im}\phi(u, v)$ as (14) and (15) and we write

$$\begin{aligned}\gamma(u,v) &= -2a_1a_2u+\gamma_{2,0}u^2+\gamma_{1,1}uv+\gamma_{0,2}v^2+\gamma_{3,0}u^3+\gamma_{2,1}u^2v+\gamma_{1,2}u^2v+\gamma_{0,3}v^3\\ &\qquad+\gamma_{4,0}u^4+\gamma_{3,1}u^3v+\gamma_{2,2}u^2v^2+\gamma_{1,3}uv^3+\gamma_{0,4}v^4+o(|u|^5+|v|^5),\\ h(u,v) &= 1+h_{1,0}u+h_{0,1}v+h_{2,0}u^2+h_{1,1}uv+h_{0,2}v^2+o(|u|^2+|v|^2).\end{aligned}$$

We already know the first coefficients of γ and h by (11) and (12). We will then compare the Taylor expansions of $\operatorname{Re}\phi(u, v)$ and $\operatorname{Im}\phi(u, v)$ obtained using (13) or using (14) and (15). Looking at all coefficients of a given order, we will get first the value of the coefficients of the Taylor expansions of γ and h of a certain order and also an equation for $\operatorname{Im}(b_1)$, $\operatorname{Im}(b_2)$ and $\operatorname{Im}(c)$.

At one point, we will have more equations than variables. These equations will have to be compatible, and will force $\Phi(z_1, z_2) = (1 - z_1z_2)/2$, which is equivalent to saying $a_1 = a_2 = \frac{1}{2}$ and $\operatorname{Im}(b_1) = \operatorname{Im}(b_2) = \operatorname{Im}(c) = 0$. This will imply the desired result.

Step 1 The goal of the first step is to show that if we have $a_1 = a_2 = \frac{1}{2}$, then we also have $\operatorname{Im}(b_1) = \operatorname{Im}(b_2) = \operatorname{Im}(c) = 0$. In addition to this, we obtain some useful equations for the following steps. [Lines 1–14]

We begin by looking at the coefficients of uv^2 in the real part of $\Phi(u, v)$. Using on the one hand (13) and on the other hand (14) we conclude that

$$\gamma_{0,2} = \frac{-6a_1^3\operatorname{Im}(b_2) - 6a_2^3\operatorname{Im}(b_1) + a_1a_2(a_1+a_2)\operatorname{Im}(c)}{8a_1a_2}.$$

We then obtain the first equation for $\operatorname{Im}(b_1)$, $\operatorname{Im}(b_2)$ and $\operatorname{Im}(c)$ by looking at the coefficients of v^3 in the real part:

$$a_2^3\operatorname{Im}(b_1) - a_1^3\operatorname{Im}(b_2) + \frac{a_1a_2(a_2-a_1)}{2}\operatorname{Im}(c) = 0. \tag{17}$$

Since we know the value of $\gamma_{0,2}$, we can get a second equation for $\mathrm{Im}(b_1)$, $\mathrm{Im}(b_2)$ and $\mathrm{Im}(c)$ by looking at the coefficients of v^2 in the imaginary part. Hence we get

$$\frac{4a_1a_2^2-3a_2^2}{4a_1}\,\mathrm{Im}(b_1)+\frac{4a_1^2a_2-3a_1^2}{4a_2}\,\mathrm{Im}(b_2)+\frac{-8a_1a_2+a_1+a_2}{8}\,\mathrm{Im}(c)=0. \tag{18}$$

By the assumptions on a_1 and a_2, we know that $2(a_1+a_2)<3$ and hence we can solve (17) and (18) with respect to $\mathrm{Im}(b_1)$ and $\mathrm{Im}(b_2)$ to obtain

$$\mathrm{Im}(b_1)=\frac{a_1(2a_2^2+2a_1a_2+a_1-2a_2)}{2a_2^2(2a_1+2a_2-3)}\,\mathrm{Im}(c),$$
$$\mathrm{Im}(b_2)=\frac{a_2(2a_1^2+2a_1a_2-2_1+a_2)}{2a_1^2(2a_1+2a_2-3)}\,\mathrm{Im}(c).$$

In particular, we may conclude that if $\mathrm{Im}(c)=0$, then we also have $\mathrm{Im}(b_1)=\mathrm{Im}(b_2)=0$. If we substitute these values into the expression for $\gamma_{0,2}$, we obtain

$$\gamma_{0,2}=\frac{-(a_1+a_2)^2}{2(2a_1+2a_2-3)}\,\mathrm{Im}(c).$$

Now, looking at the coefficient of v^4 in the real part shows that

$$(\gamma_{0,2})^2=-\frac{a_1a_2(a_1+a_2)(a_1a_2^2+a_1^2a_2-a_1^2-a_2^2+a_1a_2)}{4},$$

and this yields our first expression for $\mathrm{Im}(c)^2$,

$$\mathrm{Im}(c)^2=\frac{-a_1a_2(2a_1+2a_2-3)^2(a_1a_2^2+a_1^2a_2-a_1^2-a_2^2+a_1a_2)}{(a_1+a_2)^3}. \tag{19}$$

From (19), it is evident that if $a_1=a_2=\frac{1}{2}$, then $\mathrm{Im}(c)=0$. □

Step 2 In this step, we want to show that $a_1=a_2$. Our first goal is to compute another equation to compare with (19). [Lines 15–21]

We begin by successively looking at the coefficients of u^2v in the real part, uv in the imaginary part and uv^3 in the real part, to obtain

$$\gamma_{1,1}=\frac{a_1-a_2}{2}\,\mathrm{Im}(c),\qquad h_{0,1}=\frac{(a_1-a_2)(4a_1+4a_2-3)}{4a_1a_2(2a_1+2a_2-3)}\,\mathrm{Im}(c),$$
$$\gamma_{0,3}=\frac{-a_1+a_2}{24a_1a_2(2a_1+2a_2-3)}\times$$
$$(20{a_1}^2{a_2}^4+40{a_1}^3{a_2}^3+20{a_1}^4{a_2}^2-12a_1{a_2}^4-54{a_1}^2{a_2}^3-54{a_1}^3{a_2}^2$$
$$-12{a_1}^4a_2+18a_1{a_2}^3+18{a_1}^2{a_2}^2+18{a_1}^3a_2+3(a_1+a_2)^2\,\mathrm{Im}(c)^2).$$

Using these values, we will investigate the coefficient of v^3 in the imaginary part. This term depends indeed only on $\gamma_{1,1}$, $h_{0,1}$ and $\gamma_{0,3}$. Using the above expression, we obtain our second equation on $\mathrm{Im}(c)^2$:

$$\frac{-3(a_1-a_2)(a_1+a_2)^2(a_1+a_2-1)}{4a_1a_2(2a_1+2a_2-3)^2}\,\mathrm{Im}(c)^2 = \frac{-(a_1-a_2)(3a_1a_2^2+3a_1^2a_2-a_1^2-a_2^2-a_1a_2)}{4}. \tag{20}$$

At this stage, we have to consider several cases. Assuming that $a_1 - a_2 \neq 0$ and $a_1 + a_2 - 1 \neq 0$, we may compute

$$\mathrm{Im}(c)^2 = \frac{-a_1a_2(2a_1+2a_2-3)^2(3a_1a_2^2+3a_1^2a_2-a_1^2-a_2^2-a_1a_2)}{3(a_1+a_2)^2(a_1+a_2-1)}. \tag{21}$$

The only possibility is that the two values for $\mathrm{Im}(c)^2$ have to coincide. Equating (19) and (21) and simplifying, we obtain

$$\frac{a_1a_2(2a_1+2a_2-3)^2P(a_1,a_2)}{3(a_1+a_2)^2(a_1+a_2-1)} = 0,$$

where

$$P(a_1,a_2) = 2a_1^3+2a_2^3+a_1a_2^2+a_1^2a_2-3a_1^2-3a_2^2+3a_1a_2.$$

Since $2(a_1+a_2) < 3$, the only possibility is that $P(a_1,a_2)$ vanishes somewhere in the domain

$$\Omega = \{(a_1,a_2)\in(0,1)^2 : a_2 < a_1,\ a_2 < 1-a_1\}.$$

We first look at what happens on the boundary, where we have

$$\begin{aligned} P(a_1,0) &= 2a_1^3-3a_1^2 &&<0 \quad \text{provided } a_1\in(0,1),\\ P(a_1,a_1) &= 3a_1^2(2a_1-1) &&<0 \quad \text{provided } a_1\in(0,1/2),\\ P(a_1,1-a_1) &= -(2a_1-1) &&<0 \quad \text{provided } a_1\in(1/2,1). \end{aligned}$$

Hence, P is negative on the boundary of Ω, except at $(1/2,1/2)$. Hence, if P vanishes in Ω, then it admits a critical point there. Now, we consider the system

$$\begin{cases} 0 = \frac{\partial P}{\partial a_1}(a_1,a_2) = 6a_1^2+a_2^2+2a_1a_2-6a_1+3a_2,\\ 0 = \frac{\partial P}{\partial a_2}(a_1,a_2) = a_1^2+6a_2^2+2a_1a_2+3a_1-6a_2. \end{cases}$$

The solutions of this system are easily found to be at the intersection of two distinct ellipses. We cannot have more than two points of intersection and we have two trivial solutions, $(0,0)$ and $\left(\frac{1}{3},\frac{1}{3}\right)$. Hence, none of the critical points of P are inside Ω. Hence we get a contradiction, and we have finished this case.

Hence we must have $a_1+a_2 = 1$ or $a_1 = a_2$. Let us now investigate the case $a_1+a_2 = 1$. Looking at (20), this means either $a_1 = a_2 = \frac{1}{2}$ (and we are done) or

$$3a_1a_2^2+3a_1^2a_2-a_1^2-a_2^2-a_1a_2 = 0.$$

Taking into account that $a_2 = 1 - a_1$, we get the equation $-4a_1^2 + 4a_1 - 1 = 0$ which admits the single solution $a_1 = \frac{1}{2}$ and we get the same conclusion. Hence, the only remaining possibility is that $a_1 = a_2$. $\square$

Step 3 We are left to deal with the case $a_1 = a_2 = a \in \left(0, \frac{1}{2}\right]$. We can no longer use (21) and need to find another equation for $\operatorname{Im}(c)^2$. [Lines 21–30]

We will be looking at the coefficient before v^4 in the imaginary part. By considering $\gamma \times h$, we see that this coefficient is equal to

$$\gamma_{0,4} + \gamma_{0,3} h_{0,1} + \gamma_{0,2} h_{0,2}.$$

Hence, we are left to compute $\gamma_{0,4}$ and $h_{0,2}$. First, we compute some auxiliary values. By looking at u^3 in the real part, u^2 in the imaginary part and $u^2 v^2$ in the real part, respectively, we obtain

$$\gamma_{2,0} = -\frac{a(2a-1)}{3a-4} \operatorname{Im}(c), \qquad h_{1,0} = \frac{(2a-1)(4a-1)}{2a(4a-3)} \operatorname{Im}(c),$$
$$\gamma_{1,2} = \frac{a(2a-1)(48a^4 - 72a^3 + 27a^2 + 4\operatorname{Im}(c)^2)}{4(4a-3)^2}.$$

Knowing these values, we look at the coefficients of uv^4 in the real part and uv^2 and the imaginary part, respectively, to obtain

$$\gamma_{0,4} = -a \operatorname{Im}(c) \frac{32a^5 - 96a^4 + 90a^3 + 12a\operatorname{Im}(c)^2 - 27a^2 - 6\operatorname{Im}(c)^2}{6(4a-3)^3},$$
$$h_{0,2} = (-2a+1)\frac{128a^5 - 240a^4 + 144a^3 + 16a\operatorname{Im}(c)^2 - 27a^2 - 8\operatorname{Im}(c)^2}{8a(4a-3)^2}.$$

Finally, we investigate the coefficient of v^4 in the imaginary part to obtain the equation

$$a(2a-1)\frac{16a^4 - 32a^3 + 15a^2 + 4\operatorname{Im}(c)^2}{4(4a-3)^2} \operatorname{Im}(c) = 0.$$

Now, if $\operatorname{Im}(c) = 0$ and $a_1 = a_2 = a$, it follows at once from (19) that $a = \frac{1}{2}$, since $a \in \left(0, \frac{1}{2}\right]$. Conversely, we divide away $\operatorname{Im}(c)$ and solve for $\operatorname{Im}(c)^2$ to obtain the equation

$$\operatorname{Im}(c)^2 = -a^2 \frac{(4a-3)(4a-5)}{4}.$$

Now, this has to be equal to (19), and we find

$$-\frac{a(2a-1)(4a-3)^2}{8} = -\frac{a^2(4a-3)(4a-5)}{4}.$$

Here the only solutions are $a = 0$ and $a = \frac{3}{4}$, neither of which lie in $\left(0, \frac{1}{2}\right]$. Hence the assumption $\operatorname{Im}(c) \neq 0$ must be wrong and we conclude $a_1 = a_2 = \frac{1}{2}$. $\square$

7. Remarks and further examples

If we look more closely at the map Φ defined in Lemma 9 (the negative part of Theorem 3), then we may observe that these counterexamples all satisfy

$$\operatorname{Re}\Phi(e^{i\theta_1},\dots,e^{i\theta_d})=\tfrac{1}{2}\theta_1^2+o(\theta_1^2)\quad\text{and}\quad\operatorname{Im}\Phi(e^{i\theta_1},\dots,e^{i\theta_d})=-\theta_1+o(|\theta_1|).$$

Hence, $J(\Phi,\mathbf{1})=1$ and, using the terminology of the Section 5, we have dependence. Our next results shows that we may also have noncompactness if $J(\Phi,w)=0$ for some $w\in\mathbb{T}^d$.

Theorem 14. *There are polynomials φ with unrestricted range, of any complex dimension $d\geq 2$ and of any complex degree ≥ 4 for which the corresponding composition operator $\mathscr{C}_\varphi$ is noncompact and such that they admit a minimal Bohr lift Φ satisfying $J(\Phi,w)=0$ for any $w\in\mathbb{T}^d$ with* $\operatorname{Re}\Phi(w)=0$.

Proof. Let $\delta>0$ and define

$$\Phi(z_1,z_2)=2(1-z_1)+(1-z_1)^2(1-\delta(1-z_2)-\delta(1-z_1)(1-z_2)).$$

Let $\varphi(s)=\Phi(p_1^{-s},p_2^{-s})$. Clearly Φ is a minimal Bohr lift of φ. Then a computation shows that

$$\begin{aligned}&\operatorname{Re}\Phi(z_1,z_2)\\&=2(1-\cos x)((1-\cos x)(1+2\delta(\sin x\sin y+(1-\cos y)\cos x))+\delta(1-\cos y)).\end{aligned}$$

Clearly, for small enough $\delta>0$ this quantity is nonnegative. Hence $\varphi\in\mathscr{G}$ and $J(\Phi,(1,1))=0$ because of the relation between a_1 and b_1. Considerations similar to those of Lemma 9 show that $\mathscr{C}_\varphi$ cannot be compact. To produce a counterexample with degree 3 and a bigger complex dimension d, we may simply replace $(1-z_2)$ with

$$\frac{1}{d-1}\cdot\sum_{j=2}^{d}(1-z_j)$$

in the definition of Φ. The production of examples with degree ≥ 5 is easier. We may just set

$$\Phi(z)=(1-z_1)+\tfrac{1}{2}(1-z_1)^2+\delta(1-z_1)^4P(z),$$

where $P(z)=P(z_1,z_2,\dots,z_d)$ is any polynomial. The proof now follows that of Lemma 9. $\square$

The reason the first counterexample in Theorem 14 works is that we have a cancellation of the term $(1-\cos x)\sin x\sin y$. It seems difficult to obtain the same cancellation if we restrict ourselves to degree 3 and require $J=0$.

Question. Is it possible to construct a counterexample of degree 3 with $J=0$?

An answer to the question would in a certain sense improve the optimality of Lemma 10, but it would not yield the complete answer to which Dirichlet polynomials in $\mathscr{G}$ induce compact composition operators. Indeed, the natural next point of investigation would be this: What happens when the "quartic form" is degenerate?

In this case, terms of degree 5 also have to disappear. This follows by the mapping properties, and the argument is identical to the one used to show that degree 3 terms disappear in the case $J = 0$ given above. Hence we are reduced to studying a "sextic form".

Our counterexamples can be modified to work in this case, but they now have degree 6 and 7. Degree ≤ 3 will also easily reduce to the case of Theorem 2 in the same manner as $J = 0$ did for degree ≤ 2. However, the cases with degree 4 and 5 would need further investigation. Even if we could solve this case, we would need to investigate the case when the "sextic form" is degenerate and this leads to the "octic form" and so on.

Remark. The previous counterexample shows that we cannot deduce Theorem 2 from Theorem 3. Indeed, it is easy to construct symbols $\varphi \in \mathscr{G}$ which may be written $\varphi(s) = \sum_{j=1}^{d} P_j(p_j^{-s})$ and such that $J(\Phi, \mathbf{1}) = 0$ for Φ a minimal Bohr lift of φ. Indeed, we may consider

$$P_j(z) = (1 - z) + \tfrac{1}{2}(1 - z)^2 + \delta(1 - z)^4 Q_j(z)$$

where Q_j is an arbitrary polynomial and $\delta > 0$ is sufficiently small. Then $\mathscr{C}_\varphi$ is compact by Theorem 2 if $d \geq 2$ but this cannot be deduced from Lemma 10.

This construction can be generalized to show that Theorem 2 can handle a variety of different interesting cases not covered by Theorem 3. In fact, given any d positive integers k_j, we may find a polynomial $\Phi(z) = \sum_{j=1}^{d} \Phi_j(z_j)$ which is a minimal Bohr lift of some $\varphi \in \mathscr{G}$, with $\operatorname{Re} \Phi(z_1, \ldots, z_d) = 0$ if and only if $z = \mathbf{1}$ and here we have the expansion

$$\operatorname{Re} \Phi(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \sum_{j=1}^{d} \theta_j^{2k_j} + o\Bigl(\sum_{j=1}^{d} \theta_j^{2k_j}\Bigr),$$

$$\operatorname{Im} \Phi(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \sum_{j=1}^{d} a_1^{(j)} \theta_j + o\Bigl(\sum_{j=1}^{d} |\theta_j|\Bigr).$$

As remarked upon in the proof of Theorem 2, we must have $a_1^{(j)} > 0$. The construction of such a polynomial is immediate from our next result.

Lemma 15. *For any $k \in \mathbb{N}$, there is a polynomial $\Phi : \mathbb{C} \to \mathbb{C}$ which satisfies* $\operatorname{Re} \Phi(e^{ix}) = (1 - \cos x)^k$.

Proof. Fix N with $k \le 2N$, and for real numbers a_n and b_n consider

$$\Phi(z) = \sum_{n=1}^{N} \frac{(-1)^{n-1}}{2^n}\big(a_n(1-z)^{2n-1} - b_n(1-z)^{2n}\big).$$

Our first goal is to expand the real part of $\Phi(e^{ix})$ as a degree $2N$ polynomial in $(1-\cos x)$ with no constant term. To this end, we compute

$$(1-e^{ix})^{2n-1} = e^{\frac{i(2n-1)x}{2}}(e^{-\frac{ix}{2}} - e^{\frac{ix}{2}})^{2n-1} = e^{\frac{i(2n-1)x}{2}} 2^{2n-1}(-1)^n i \sin^{2n-1}\Big(\frac{x}{2}\Big).$$

We use $2\sin^2(x/2) = 1-\cos x$, and obtain

$$\begin{aligned} \operatorname{Re}(1-e^{ix})^{2n-1} &= 2^{2n-1}(-1)^{n-1}\sin^{2n-1}\Big(\frac{x}{2}\Big)\sin\Big(nx-\frac{x}{2}\Big) \\ &= (-1)^{n-1}2^n(1-\cos x)^n \frac{\sin\big(nx-\frac{x}{2}\big)}{2\sin\big(\frac{x}{2}\big)}. \end{aligned} \tag{22}$$

Similarly, we obtain

$$\operatorname{Re}(1-e^{ix})^{2n} = 2^{2n}(-1)^n\sin^{2n}\Big(\frac{x}{2}\Big)\cos(nx) = (-1)^n 2^n(1-\cos x)^n\cos(nx). \tag{23}$$

To continue the computations, we introduce the Chebyshev polynomials

$$U_n(y) = \sum_{j=0}^{n}(-2)^j\frac{(n+j+1)!}{(n-j)!(2j+1)!}(1-y)^j, \quad T_n(y) = n\sum_{j=0}^{n}(-2)^j\frac{(n+j-1)!}{(n-j)!(2j)!}(1-y)^j.$$

These polynomials are relevant due to the formulas $\sin nx = \sin(x)U_{n-1}(\cos x)$ and $\cos nx = T_n(\cos x)$. We record the following coefficients:

$$u_{n-2}^{(n)} = (-2)^{n-1}\Big(-\frac{(2n-1)(n-2)}{2}\Big), \qquad u_{n-1}^{(n)} = (-2)^{n-1}(2n),$$
$$u_n^{(n)} = (-2)^{n-1}(-2), \qquad t_{n-1}^{(n)} = (-2)^{n-1}(n), \qquad t_n^{(n)} = (-2)^{n-1}(-1).$$

Now, we rewrite (23) as

$$\operatorname{Re}(1-e^{ix})^{2n} = (-1)^n 2^n(1-\cos x)^n T_n(\cos x),$$

which is then clearly a degree $2n$ polynomial in $(1-\cos x)$ with no constant term. For (22) we have to work a bit more, so we first compute

$$\frac{\sin\big(nx-\frac{x}{2}\big)}{2\sin\big(\frac{x}{2}\big)} = \frac{\sin(nx)\cos\big(\frac{x}{2}\big) - \cos(nx)\sin\big(\frac{x}{2}\big)}{2\sin\big(\frac{x}{2}\big)} = \cos^2\Big(\frac{x}{2}\Big)U_{n-1}(\cos x) - \frac{T_n(\cos x)}{2},$$

which implies that we may rewrite (22) as

$$\operatorname{Re}(1-e^{ix})^{2n-1} = (-1)^{n-1}2^n(1-\cos x)^n\Big(\Big(1-\frac{1-\cos x}{2}\Big)U_{n-1}(\cos x) - \frac{T_n(\cos x)}{2}\Big).$$

Again we observe that this is a polynomial of degree $2n$ in $(1-\cos x)$ with no constant term. In total, we have

$$\operatorname{Re}\Phi(1-e^{ix}) = \sum_{m=1}^{2N} c_m(1-\cos x)^m = \sum_{n=1}^{N}(a_n P_n(1-\cos x) + b_n Q_n(1-\cos x)),$$

where

$$P_n(y) = \sum_{j=0}^{n} d_j^{(n)} y^{n+j} = y^n\Big(\Big(1-\frac{y}{2}\Big)U_{n-1}(1-y) - \frac{T_n(1-y)}{2}\Big),$$
$$Q_n(y) = \sum_{j=0}^{n} e_j^{(n)} y^{n+j} = y^n T_n(1-y).$$

Given any choice of c_m (for instance $c_m = 0$ for $m \neq k$ and $c_k = 1$), we now have $2N$ linear equations and $2N$ unknowns, a_n and b_n for $1 \leq n \leq N$. We will now show that this system can always be solved.

We first observe that a_n and b_n only have an effect on c_m when $n \leq m \leq 2n$. Ordering the unknowns as $a_N, b_N, a_{N-1}, b_{N-1}, \ldots, a_1, b_1$ and the data as $c_{2N}, c_{2N-1}, \ldots, c_1$, this means that the matrix of our system can be written in upper triangular block form, where the blocks on the diagonal are

$$\begin{pmatrix} d_{n-1}^{(n)} & e_{n-1}^{(n)} \\ d_n^{(n)} & e_n^{(n)} \end{pmatrix}, \qquad n = N, N-1, \ldots, 1.$$

We know that $e_{n-1}^{(n)} = t_{n-1}^{(n)}$ and $e_n^{(n)} = t_n^{(n)}$, which we recorded above. It is now easy to verify that

$$d_{n-1}^{(n)} = u_{n-1}^{(n)} - \frac{u_{n-2}^{(n)}}{2} - \frac{t_{n-1}^{(n)}}{2} = (-2)^{n-1}\Big(\frac{3n}{2} + \frac{(2n-1)(n-2)}{4}\Big),$$
$$d_n^{(n)} = u_n^{(n)} - \frac{u_{n-1}^{(n)}}{2} - \frac{t_n^{(n)}}{2} = (-2)^{n-1}\Big(-\frac{3}{2} - n\Big).$$

Hence we are reduced to considering the equation

$$0 = \frac{d_{n-1}^{(n)} e_n^{(n)} - d_n^{(n)} e_{n-1}^{(n)}}{4^{n-1}} = \Big(\frac{3n}{2} + \frac{(2n-1)(n-2)}{4}\Big)(-1) - \Big(-\frac{3}{2} - n\Big)n$$
$$= n^2 - \frac{(2n-1)(n-2)}{4},$$

which has no integer solutions, and we are done. □

The construction of Φ with specific expansion facilitated by Lemma 15 will be used in the next section to prove Corollary 4.

8. Approximation numbers

In this section, we consider only the case $c_0 = 0$. We intend to estimate the decay of $a_n(\mathscr{C}_\varphi)$ for maps φ which are, in a certain sense, regular at their boundary points. For this we need as previously a careful inspection of the behavior of the Bohr lift Φ near these boundary points.

Definition. Suppose that $\varphi(s) = c_1 + \sum_{n=2}^{N} c_n n^{-s} \in \mathscr{G}$, and that φ has complex dimension d and unrestricted range. Let Φ be a minimal Bohr lift of φ and let $w \in \mathbb{T}^d$ be such that $\operatorname{Re}\Phi(w) = 0$. We say that φ is *boundary regular* at w if there exist independent linear forms $\ell_1, \dots, \ell_d$ on $\mathbb{C}^d$, even integers $k_1 \geq k_2 \geq \cdots \geq k_d$ and real numbers $b_1, \dots, b_d, \tau$ with $b_1 \neq 0$ such that

$$\operatorname{Re}\Phi(e^{i\theta_1}w_1, \dots, e^{i\theta_d}w_d) = \ell_1(\theta)^{k_1} + \cdots + \ell_d(\theta)^{k_d} + \sum_{j=1}^{d} o\big(\ell_j^{k_j}(\theta)\big) \tag{24}$$

$$\operatorname{Im}\Phi(e^{i\theta_1}w_1, \dots, e^{i\theta_d}w_d) = \tau + b_1\ell_1(\theta) + \cdots + b_d\ell_d(\theta) + o\bigg(\sum_{j=1}^{d}|\ell_j(\theta)|\bigg). \tag{25}$$

We define the *compactness index* of φ at w as

$$\eta_{\varphi,w} = \bigg(\sum_{j=2}^{d}\frac{1}{k_j}\bigg) \times \frac{k_1}{2(k_1-1)}.$$

If every boundary point is boundary regular, we say that φ is *boundary regular*.

The proof of Theorem 2 then shows that given a boundary regular map φ, the composition operator $\mathscr{C}_\varphi$ is compact if and only if $d \geq 2$. We shall now assume that there is only one point $w \in \mathbb{T}^d$ such that $\operatorname{Re}\Phi(z) = 0$. In this case, we let the *compactness index* of φ be $\eta_\varphi := \eta_{\varphi,w}$.

The main theorem of this section now reads:

Theorem 16. *Let $\varphi(s) = c_1 + \sum_{n=2}^{N} c_n n^{-s} \in \mathscr{G}$ have unrestricted range and complex dimension d. Let Φ be a minimal Bohr lift and assume that there exists a unique $w \in \mathbb{T}^d$ such that* $\operatorname{Re}\Phi(w) = 0$. *Suppose moreover that φ is boundary regular at w. Then*

$$\left(\frac{1}{n}\right)^{\eta_\varphi} \ll a_n(\mathscr{C}_\varphi) \ll \left(\frac{\log n}{n}\right)^{\eta_\varphi}.$$

This statement may be applied to several cases:

Corollary 17. *Let $\varphi(s) = c_1 + \sum_{n=2}^{N} c_n n^{-s} \in \mathscr{G}$ have unrestricted range and complex dimension d. Let Φ be a minimal Bohr lift of φ and assume that there exists a unique $w \in \mathbb{T}^d$ such that* $\operatorname{Re}\Phi(w) = 0$ *and that $J(\Phi, w) = d$. Then*

$$\left(\frac{1}{n}\right)^{(d-1)/2} \ll a_n(\mathscr{C}_\varphi) \ll \left(\frac{\log n}{n}\right)^{(d-1)/2}.$$

Proof. With these assumptions, φ is boundary regular at w with $k_1 = \cdots = k_d = 2$. $\square$

In particular, this corollary covers the result of Queffélec and Seip for linear symbols (3), as well as the map φ_1 given in (4). We may also apply Theorem 16 to the maps considered in Theorem 2. In this case, one has simply $\ell_j(\theta) = \theta_j$ (up to a reordering of the terms).

Another interesting application of Theorem 16 is that we may distinguish the Schatten classes of bounded linear operators on $\mathscr{H}^2$ using composition operators, as mentioned in the introduction.

Proof of Corollary 4. Let $p' < q'$ and $\varepsilon > 0$ be such that $p \le p'/2$ and $\left(\frac{1}{2}+\varepsilon\right)q' \le q$. Then let $d \ge 2$ and $k \ge 2$ even such that

$$p' < \frac{d-1}{k} < q' \quad \text{and} \quad \frac{1}{2} < \frac{k}{2(k-1)} < \frac{1}{2} + \varepsilon.$$

By Lemma 15, we know that there exists a boundary regular polynomial $\Phi : \mathbb{T}^d \to \mathbb{C}_0$ such that $\operatorname{Re}\Phi(w) = 0$ if and only if $w = \mathbf{1}$ and

$$\Phi(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \theta_1^k + \cdots + \theta_d^k + o(\theta_1^k) + \cdots + o(\theta_d^k).$$

Letting $\varphi \in \mathscr{G}$ any map such that Φ is a minimal Bohr lift of φ, we immediately get

$$\left(\frac{1}{n}\right)^{\frac{d-1}{k} \times \frac{k}{2(k-1)}} \ll a_n(\mathscr{C}_\varphi) \ll \left(\frac{\log n}{n}\right)^{\frac{d-1}{k} \times \frac{k}{2(k-1)}},$$

which completes the proof. $\square$

Theorem 16 may also be applied to many other maps. We will consider here the map φ_2 given in (4). Its boundary regularity is different than that of φ_1, and hence the degree of compactness is also different.

Example. Let $\varphi_2(s) = \frac{13}{2} - 4 \cdot 2^{-s} - 4 \cdot 3^{-s} + 2 \cdot 6^{-s}$ as in (4) and let Φ be its minimal Bohr lift. It can be shown that $\operatorname{Re}\Phi(w) = 0$ for $w \in \mathbb{T}^2$ if and only if $w = (1, 1)$, and

$$\begin{aligned} \operatorname{Re}\Phi(e^{i\theta_1}, e^{i\theta_2}) &= \ell_1(\theta)^4 + \ell_2(\theta)^2 + o(\ell_1^4(\theta)) + o(\ell_2^2(\theta)), \\ \operatorname{Im}\Phi(e^{i\theta_1}, e^{i\theta_2}) &= -2\ell_1(\theta) + o(|\ell_1(\theta)| + |\ell_2(\theta)|), \end{aligned}$$

where $\ell_1(\theta) = \theta_1 + \theta_2$ and $\ell_2(\theta) = \theta_1 - \theta_2$. Hence $\eta_{\varphi_2} = \left(\frac{1}{2}\right) \times \left(\frac{4}{6}\right) = \frac{1}{3}$.

The remaining part of this section is devoted to the proof of Theorem 16. We use the scheme introduced by Queffélec and Seip in [2015a] in the context of Dirichlet series (see also [Queffélec and Seip 2015b] for similar works on the classical Hardy space of the disk). Their method is based on Carleson measures, interpolation sequences and model spaces. In Section 8.1, we survey these tools and give a couple of lemmas. Section 8.2 is devoted to the proof of the upper bound, in a more general context, whereas Section 8.3 will be devoted to the lower bound.

8.1. *Tools.*

The Hyperbolic Metric. The pseudohyperbolic metric on the half-plane $\mathbb{C}_0$ is defined by

$$\rho(z, w) = \left|\frac{z-w}{z+\overline{w}}\right| = \frac{1-e^{-d(z,w)}}{1+e^{d(z,w)}},$$

where $d(z, w)$ is the hyperbolic distance between two points z and w in $\mathbb{C}_0$. The hyperbolic length of a curve $\Gamma \subset \mathbb{C}_0$ is given by the integral

$$L_p(\Gamma) = \int_\Gamma \frac{|dz|}{\operatorname{Re} z}.$$

Carleson Measures and Interpolating Sequences. Let H be a Hilbert space of functions defined on some measurable set Ω in $\mathbb{C}$. A nonnegative Borel measure μ on Ω is a *Carleson measure* for H if there exists some constant $C > 0$ such that

$$\int_\Omega |f(z)|^2 \, d\mu(z) \leq C\|f\|_H^2,$$

for every f in H. The smallest possible C will be called the *Carleson norm* of μ with respect to H and will be denoted by $\|\mu\|_{\mathscr{C},H}$.

We also assume that the linear point evaluation is bounded at any $z \in \Omega$. Then H admits a reproducing kernel $K_z^H \in H$ for any $z \in \Omega$ which satisfies $f(z) = \langle f, K_z^H\rangle$ for every $f \in H$. We then say that a sequence $Z = (z_m)$ of distinct points in Ω is a *Carleson sequence* for H if the measure

$$\mu_{Z,H} := \sum_m \left\| K_{z_m}^H \right\|_H^{-2} \delta_{z_m}$$

is a *Carleson measure* for H.

We say that a sequence $Z = (z_m)$ of distinct points in Ω is an *interpolating sequence* for H if the interpolation problem $f(z_m) = a_m$ has a solution $f \in H$ whenever the admissibility condition

$$\sum_m |a_m|^2 \left\| K_{z_m}^H \right\|_H^{-2} < \infty$$

is satisfied. By the open mapping theorem, if Z is an interpolating sequence for H, there is a constant $C > 0$ such that we can solve $f(z_m) = a_m$ with f satisfying

$$\|f\|_H \leq C\left(\sum_m |a_m|^2 \left\| K_{z_m}^H \right\|_H^{-2}\right)^{1/2}.$$

The smallest constant C with this property will be called the *constant of interpolation* of Z and will be denoted by $M_H(Z)$.

We shall consider the two spaces $H = \mathscr{H}^2$ and $H = H^2(\mathbb{T}^d)$. Then we have, respectively, $\Omega = \mathbb{C}_{1/2}$ and $\Omega = \mathbb{D}^d$, and moreover

$$\big\| K_s^{\mathscr{H}^2} \big\|^{-2} = [\zeta(2\operatorname{Re} s)]^{-1} \quad \text{and} \quad \big\| K_z^{H^2(\mathbb{T}^d)} \big\|^{-2} = \prod_{j=1}^{d} (1 - |z_j|^2).$$

We will need the three following lemmas.

Lemma 18. *Let μ be a Borel measure on $\overline{\mathbb{C}_0}$, let $\sigma \in (0, 1)$ and $R > 0$. Assume that μ is supported on the rectangle $0 \le \operatorname{Re} s \le \sigma$, $|\operatorname{Im} s| \le R$. Then*

$$\|\mu\|_{\mathscr{C},\mathscr{H}^2} \ll_R \sup_{\varepsilon>0,\, \tau\in\mathbb{R}} \frac{\mu(Q(\tau,\varepsilon))}{\varepsilon} \le 2 \sup_{\varepsilon\in(0,\sigma),\, \tau\in\mathbb{R}} \frac{\mu(Q(\tau,\varepsilon))}{\varepsilon}.$$

Proof. The first inequality is Lemma 2.3 in [Queffélec and Seip 2015a] (the involved constant does not depend on $\sigma \in (0, 1)$). The second follows from the inequality

$$\sup_{\tau\in\mathbb{R}} \frac{\mu(Q(\tau, 2^{k+1}\sigma))}{2^{k+1}\sigma} \le \sup_{\tau\in\mathbb{R}} \frac{\mu(Q(\tau, 2^{k}\sigma))}{2^{k}\sigma},$$

valid for any $k \ge 0$. Indeed, for any $\tau \in \mathbb{R}$ and any $k \ge 0$, we may find $\tau_1, \tau_2 \in \mathbb{R}$ such that

$$\mu(Q(\tau, 2^{k+1}\sigma)) = \mu(Q(\tau_1, 2^k\sigma)) + \mu(Q(\tau_2, 2^k\sigma)),$$

since the support of μ is contained in $0 \le \operatorname{Re} s \le \sigma$. □

Lemma 19. *Let $\nu > 0$. There exists $C > 0$ such that, for any $\delta \in (0, 1/\nu)$, $M_{\mathscr{H}^2}(S_\delta) \le C$, where $S_\delta = (s_m)_{m=1}^{1/\delta}$ with $s_m = \frac{1}{2} + \nu\delta + im\delta$.*

Proof. The proof can be found in [Queffélec and Seip 2015a, § 8.2]. □

Lemma 20. *Let $C_1, C_2 > 0$. There exists $D > 0$ such that for any $\delta > 0$ and any (finite) sequence*

$$Z = (Z(\alpha)) = ((1 - \rho_1(\alpha))e^{i\theta_1(\alpha)}, \dots, (1 - \rho_d(\alpha))e^{i\theta_d(\alpha)})$$

in $\mathbb{D}^d$ satisfying

- $\sup_{j=1,\dots,d} |\theta_j(\alpha) - \theta_j(\beta)| \ge C_1\delta$, *when $\alpha \ne \beta$,*
- $\rho_j(\alpha) \le C_2\delta$, *for any α and $j = 1, \dots, d$,*

we have $\big\|\mu_{Z,H^2(\mathbb{T}^d)}\big\|_{\mathscr{C},H^2(\mathbb{T}^d)} \le D$.

Proof. To each point $Z(\alpha)$, we associate a rectangle R_α on the distinguish boundary $\mathbb{T}^d$ centered at

$$\Big(\frac{z_1(\alpha)}{|z_1(\alpha)|}, \dots, \frac{z_d(\alpha)}{|z_d(\alpha)|}\Big),$$

with side lengths $2(1-|z_1(\alpha)|),\dots,2(1-|z_d(\alpha)|)$. By Chang's characterization of Carleson measures on the polydisc (see [Berndtsson et al. 1987] or [Chang 1979]), it is enough to show that we for all open sets $\mathcal{U}$ of $\mathbb{T}^d$ have

$$\sum_{R_\alpha\subset\mathcal{U}} \boldsymbol{m}_d(R_\alpha) \le D\boldsymbol{m}_d(\mathcal{U}).$$

If R is some rectangle in $\mathbb{T}^d$ and $\lambda>0$, denote by λR the rectangle with the same center and side lengths multiplied by λ. Then our assumptions on Z imply that there exists some $\lambda\in(0,1)$ depending only on C_1 and C_2 such that the rectangles R_α are pairwise disjoint. Thus

$$\sum_{R_\alpha\subset\mathcal{U}} \boldsymbol{m}_d(\mathcal{U}) \le \sum_{R_\alpha\subset\mathcal{U}} \frac{1}{\lambda^d}\boldsymbol{m}_d(\lambda R_\alpha) \le \frac{1}{\lambda^d}\boldsymbol{m}_d\biggl(\bigcup_{R_\alpha\subset\mathcal{U}} \lambda R_\alpha\biggr) \le \frac{1}{\lambda^d}\boldsymbol{m}_d(\mathcal{U}),$$

which completes the proof with $D=1/\lambda^d$. □

The Queffélec–Seip Method. We have to introduce additional conventions. For $\varphi\in\mathcal{G}$ and Ω a compact subset of $\mathbb{C}_0$, we denote by $\mu_{\varphi,\Omega}$ the nonnegative Borel measure on $\overline{\mathbb{C}_0}$ defined by

$$\mu_{\varphi,\Omega}(E) := \boldsymbol{m}_\infty(\{z\in\mathbb{T}^\infty : \Phi(z)\in E\setminus\Omega\}).$$

Next, assume that φ has complex dimension d and Bohr lift $\Phi:\mathbb{C}^d\to\mathbb{C}$. Let $S=(s_m)$ be a sequence of n points in $\mathbb{C}_{1/2}$ and let Z be a finite sequence of points in $\mathbb{D}^d$ such that $\Phi(Z)=S-\frac{1}{2}$. We set

$$N_\Phi(s_m;Z) := \sum_{z\in Z\cap\Phi^{-1}(s_m-1/2)} \bigl\| K_z^{H^2(\mathbb{T}^d)}\bigr\|^{-2}.$$

We state Theorem 4.1 of [Queffélec and Seip 2015a] as the following lemma (we have modified it slightly to take into account our normalization):

Lemma 21. *Let $\varphi(s)=\sum_{n=1}^\infty c_n n^{-s}\in\mathcal{G}$ such that $\varphi(\mathbb{C}_0)$ is bounded.*

(a) *Let $\sigma>0$ and Ω be a compact subset of $\overline{\mathbb{C}_\sigma}$. Let B be a Blaschke product on $\mathbb{C}_0$ of degree n whose zeros lie in Ω. Then*

$$a_n(\mathscr{C}_\varphi) \le \biggl(\sup_{s\in\Omega}|B(s)|^2\zeta(1+2\sigma) + \sup_{\varepsilon>0,\tau\in\mathbb{R}} \frac{\mu_{\varphi,\Omega}(Q(\tau,\varepsilon))}{\varepsilon}\biggr)^{1/2}.$$

(b) *Assume that φ has complex dimension d. Let S and Z be finite sets in respectively $\mathbb{C}_{1/2}$ and $\mathbb{D}^d$ such that $\Phi(Z)=S-\frac{1}{2}$. Then*

$$a_n(\mathscr{C}_\varphi) \ge [M_{\mathscr{H}^2}(S)]^{-1} \bigl\|\mu_{Z,H^2(\mathbb{T}^d)}\bigr\|^{-1/2}_{\mathscr{C},H^2(\mathbb{T}^d)} \inf_m [N_\Phi(s_m;Z)\zeta(2\operatorname{Re}s_m)]^{1/2}.$$

8.2. ***The Upper Bound.*** Let $\varphi(s) = \sum_{n=1}^{\infty} c_n n^{-s} \in \mathscr{G}$ and suppose that $\varphi(\mathbb{C}_0)$ bounded. By Lemma 5, $\mathscr{C}_\varphi$ is compact if and only if $\mu_\varphi(Q(\tau, \varepsilon)) = o(\varepsilon)$ uniformly in $\tau \in \mathbb{R}$. We are planning to get an upper bound of $a_n(\mathscr{C}_\varphi)$ depending on the behavior of $\sup_{\tau \in \mathbb{R}} \mu_\varphi(Q(\tau, \varepsilon))$ with respect to ε and on the size of the image of φ near a boundary point.

Thus, let Φ be a Bohr lift of φ. We define κ_φ as the infimum of those $\kappa \geq 1$ such that there exists a constant $C > 0$ such that, for every $\tau \in \mathbb{R}$ and every $\varepsilon > 0$,

$$\boldsymbol{m}_\infty(\{z \in \mathbb{T}^\infty : \Phi(z) \in Q(\tau, \varepsilon)\}) \leq C\varepsilon^\kappa.$$

Assume now that there exists a unique $w \in \mathbb{T}^\infty$ such that $\operatorname{Re}\Phi(w) = 0$ and write $\Phi(w) = i\tau$. Let ω_φ be the infimum of the positive ω such that, for any $s \in \mathbb{C}_0$,

$$|\operatorname{Im}\varphi(s) - \tau|^\omega \leq C\big(\operatorname{Re}\varphi(s) - \tfrac{1}{2}\big).$$

Theorem 22. *Let $\varphi(s) = \sum_{n=1}^{\infty} c_n n^{-s} \in \mathscr{G}$ with $\varphi(\mathbb{C}_0)$ bounded, let Φ be a Bohr lift of φ and assume that there is a unique $w \in \mathbb{T}^\infty$ such that* $\operatorname{Re}\Phi(w) = 0$. *Then*

$$a_n(\mathscr{C}_\varphi) \ll \begin{cases} \exp(-\lambda n^{-1/2}) & \text{if } \omega_\varphi \leq 1, \\ \Big(\dfrac{\log n}{n}\Big)^{(\kappa_\varphi - 1)\times\frac{\omega_\varphi}{2(\omega_\varphi - 1)}} & \text{if } \omega_\varphi > 1. \end{cases}$$

Here λ is some positive constant depending on φ.

This theorem illustrates the following general principle for composition operators (valid beyond $\mathscr{H}^2$): the more restricted the image of the symbol is, the more compact the associated composition operator is. In particular, the case $\omega_\varphi = 1$ (the range of φ is contained in an angle) is reminiscent of [Queffélec and Seip 2015b, Theorem 1.2], where a similar result was obtained for composition operators on $H^2(\mathbb{D})$.

Before we embark upon the proof of Theorem 22, we first employ it to deduce the upper bound of Theorem 16.

Final part in the proof of the upper bound of Theorem 16. Suppose that $\varphi \in \mathscr{G}$ is a boundary regular Dirichlet polynomial, and assume that $\operatorname{Re}\Phi(\mathbf{1}) = 0$. We write

$$\operatorname{Re}\Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) = \ell_1(\theta)^{k_1} + \dots + \ell_d(\theta)^{k_d} + \sum_{j=1}^{d} o(\ell_j^{k_j}(\theta)),$$

$$\operatorname{Im}\Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) = \tau + b_1\ell_1(\theta) + \dots + b_d\ell_d(\theta) + o\bigg(\sum_{j=1}^{d} |\ell_j(\theta)|\bigg),$$

with $k_1 \geq \dots \geq k_d$ and $b_1 \neq 0$. The proof of Theorem 2 shows that we have $\kappa_\varphi \geq 1 + \sum_{j=2}^{d} 1/k_j$.

Now, let us write the Taylor expansion of $\operatorname{Re}\Phi$ and $\operatorname{Im}\Phi$ near $\mathbf{1}$, but also now for a point belonging to the unit polydisc. Writing

$$\Phi(z)=\sum_{j=1}^{d}a_j(1-z_j)+o\left(\sum_{j=1}^{d}|1-z_j|\right)$$

and $z=((1-\rho_1)e^{i\theta_1},\dots,(1-\rho_d)e^{i\theta_d})$, it is easy to get

$$\operatorname{Re}\Phi(z)=a_1\rho_1+\cdots+a_d\rho_d+\ell_1(\theta)^{k_1}+\cdots+\ell_d(\theta)^{k_d}+o\left(\sum_{j=1}^{d}\bigl(\rho_j+\ell_j^{k_j}(\theta)\bigr)\right),$$

$$\operatorname{Im}\Phi(z)=\tau+b_1\ell_1(\theta)+\cdots+b_d\ell_d(\theta)+o\left(\sum_{j=1}^{d}|\ell_j(\theta)|\right).$$

Recalling that $a_j\geq 0$ for $j=1,\dots,d$, it is easy to conclude that there exists a neighborhood $\mathcal{U}\ni\mathbf{1}$ in $\overline{\mathbb{D}}^d$ and $C>0$ such that, for all $z\in\mathcal{U}$,

$$|\operatorname{Im}\Phi(z)-\tau|^{k_1}\leq C\operatorname{Re}\Phi(z).$$

Outside $\mathcal{U}$, $\operatorname{Re}\Phi(z)$ is bounded away from 0, and $|\operatorname{Im}\Phi(z)-\tau|$ is here trivially majorized. Hence, the upper bound of Theorem 16 follows from Theorem 22. $\square$

Let us now turn to the proof of Theorem 22. The proof will be preceded by two lemmas. The first one is inspired by Lemma 3.1 in [Queffélec and Seip 2015b].

Lemma 23. *Let Ω be a bounded domain in $\mathbb{C}_0$ whose boundary is a piecewise regular Jordan curve Γ, with $L_p(\Gamma)\geq 1$. Let $s_1,\dots,s_n$ be points in Γ such that the hyperbolic length of the curve between any two points s_j and s_{j+1} is equal to $L_p(\Gamma)/n$, $1\leq j\leq n$, where $s_{n+1}=s_1$. Let B be the Blaschke product of degree n whose zeros are precisely $s_1,\dots,s_n$. Then, for any $s\in\Omega$,*

$$|B(s)|\leq\exp\left(-C\frac{n}{L_p(\Gamma)}\right).$$

Proof. By the maximum principle, it is sufficient to prove this inequality for $s\in\Gamma$. In this case, we know that there exists some $j\in\{1,\dots,n\}$ such that $d(s,s_j)\leq L_p(\Gamma)/n$, from which we deduce that

$$d(s,s_k)\leq\frac{L_p(\Gamma)}{n}(1+|k-j|)$$

for any $k=1,\dots,n$. Using the link between the pseudo hyperbolic distance and the hyperbolic distance, we deduce that

$$|B(s)|\leq\prod_{j=1}^{n}\left(\frac{1-e^{-j\frac{L_p(\Gamma)}{n}}}{1+e^{-j\frac{L_p(\Gamma)}{n}}}\right).$$

By a Riemann sum argument, this means that

$$|B(s)| \leq \exp\left(-n\int_0^1 \ln\left(\frac{1-e^{-xL_p(\Gamma)}}{1+e^{xL_p(\Gamma)}}\right)dx\right) \leq \exp\left(-\frac{n}{L_p(\Gamma)}\int_{e^{-L_p(\Gamma)}}^1 \frac{1}{y}\ln\left(\frac{1+y}{1-y}\right)dy\right),$$

and we get the desired conclusion, since by assumption $L_p(\Gamma) \geq 1$. □

Hence, we require estimates of the hyperbolic length of some curves which are linked to the way that φ touches the boundary. Such estimates are contained in the following result:

Lemma 24. *Let $\omega \geq 1$, $\sigma \in (0, 1/2)$ and $C > 1$. Consider*

$$\Omega_{\omega,\sigma,C} = \{s \in \mathbb{C}_0 : |\operatorname{Im} s|^{\omega} \leq C \operatorname{Re}(s),\ \sigma \leq \operatorname{Re} s \leq C\}.$$

Let $\Gamma_{\omega,\sigma,C}$ denote the boundary of $\Omega_{\omega,\sigma,C}$. Then

$$L_p(\Gamma_{\omega,\sigma,C}) \ll_{\omega,C} \begin{cases} \left(\frac{1}{\sigma}\right)^{\frac{\omega-1}{\omega}} & \textit{if } \omega > 1, \\ -\ln(\sigma) & \textit{if } \omega = 1. \end{cases}$$

Proof. Consider the curves

$$\begin{aligned} \Gamma_1 &= \{s \in \mathbb{C}_0 : \qquad \operatorname{Re} s = \sigma,\ |\operatorname{Im} s| \leq C^{1/\omega}(\operatorname{Re} s)^{1/\omega}\}, \\ \Gamma_2 &= \{s \in \mathbb{C}_0 : \qquad \operatorname{Re} s = C,\ |\operatorname{Im} s| \leq C^{1/\omega}(\operatorname{Re} s)^{1/\omega}\}, \\ \Gamma_3 &= \{s \in \mathbb{C}_0 : \sigma \leq \operatorname{Re} s \leq C,\ |\operatorname{Im} s| = C^{1/\omega}(\operatorname{Re} s)^{1/\omega}\}. \end{aligned}$$

Clearly, $\Gamma_{\omega,\sigma,C} \subset \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ and it is sufficient to prove the corresponding inequalities for Γ_j, $j = 1, 2, 3$. Firstly, $L_p(\Gamma_2) \ll_{\omega,C} 1$. Regarding Γ_1,

$$L_p(\Gamma_1) = \int_{-C^{1/\omega}\sigma^{1/\omega}}^{C^{1/\omega}\sigma^{1/\omega}} \frac{dy}{\sigma} \ll_{\omega,C} \left(\frac{1}{\sigma}\right)^{\frac{\omega-1}{\omega}}$$

which is even a stronger inequality than required when $\omega = 1$. Finally,

$$L_p(\Gamma_3) \ll_{\omega,C} \int_{\sigma}^{C} \frac{\sqrt{1+x^{\frac{2}{\omega}-1}}}{x}\,dx \ll_{\omega,C} \begin{cases} -\ln(\sigma) & \text{if } \omega = 1, \\ \left(\frac{1}{\sigma}\right)^{\frac{\omega-1}{\omega}} & \text{if } \omega > 1. \end{cases}$$

The last estimate follows by inspecting the integrand near $x = 0$, since $\sigma \in (0, 1)$. □

Proof of Theorem 22. Let $\sigma \in (0, 1)$ and $n \geq 1$. Without loss of generality, we may assume that $\Phi(\mathbf{1}) = 0$. Keeping the notations of Lemma 24, there exists $C > 0$ such that

$$\varphi(\mathbb{C}_0) - \tfrac{1}{2} \subset \{s \in \mathbb{C}_0 : 0 \leq \operatorname{Re} s \leq \sigma\} \cup \Omega_{\omega_\varphi,\sigma,C}.$$

Let B be a Blaschke product of degree n defined as in Lemma 23 with $\Omega_{\omega_\varphi,\sigma,C}$. Enlarging C if necessary, we may always assume that $L_p(\Gamma_{\omega_\varphi,\sigma,C}) \geq 1$, so that the assumptions of Lemma 23 are satisfied. The set

$$\Omega = \overline{\left\{\varphi(s) - \tfrac{1}{2} : \operatorname{Re}\varphi(s) \geq \tfrac{1}{2} + \sigma\right\}}$$

is a compact subset of $\mathbb{C}_0$, and we may apply part (a) of Lemma 21. Since $\Omega \subset \Omega_{\omega_\varphi,\sigma,C}$, we obtain

$$\sup_{s\in\Omega} |B(s)|^2 \leq \exp\Big(-2C' \frac{n}{L_p(\Gamma_{\omega_\varphi,\sigma,C})}\Big).$$

Moreover, $\zeta(1+2\sigma) \ll 1/\sigma$. Finally, using Lemma 18, we obtain

$$\|\mu_{\varphi,\Omega}\|_{\mathscr{C},\mathscr{H}^2} \ll \sup_{\varepsilon\in(0,\sigma),\, \tau\in\mathbb{R}} \frac{\mu_{\varphi,\Omega}(Q(\tau,\varepsilon))}{\varepsilon} \ll \sigma^{\kappa_\varphi - 1}.$$

We will now optimize the choice of σ with respect to n. When $\omega_\varphi > 1$, we set

$$\sigma = \rho\Big(\frac{\log n}{n}\Big)^{\frac{\omega_\varphi}{\omega_\varphi - 1}},$$

where ρ is some numerical parameter to be chosen later. Then

$$\sup_{s\in\Omega} |B(s)|^2 \zeta(1+2\sigma) \leq \exp\big(-2C'\rho^{\frac{\omega_\varphi - 1}{\omega_\varphi}} \log n\big) \cdot \frac{1}{\sigma} \ll \Big(\frac{\log n}{n}\Big)^{\frac{\omega_\varphi}{\omega_\varphi - 1}(\kappa_\varphi - 1)},$$

provided $\rho > 0$ is sufficiently large. When $\omega_\varphi \leq 1$, we set $\sigma = \exp(-\rho n^{-1/2})$, so that

$$\sup_{s\in\Omega} |B(s)|^2 \zeta(1+2\sigma) \leq \exp\Big(-\frac{C''}{\rho} n^{1/2} + \rho n^{1/2}\Big),$$

and the result is proved provided $\rho > 0$ is sufficiently small. □

Remark. Our method of proof also shows, provided $\varphi(\mathbb{C}_0)$ is bounded and $\kappa_\varphi > 1$, that

$$a_n(\mathscr{C}_\varphi) \leq \Big(\frac{\log n}{n}\Big)^{\frac{\kappa_\varphi - 1}{2}}.$$

Indeed, we apply the same method with $\Omega_{\sigma,C} = \{s \in \mathbb{C}_0 : \sigma \leq \operatorname{Re} s \leq C,\ |\operatorname{Im} s| \leq C\}$ which satisfies $L_p(\Gamma) \ll_C \sigma^{-1}$. The rest of the proof remains unchanged.

8.3. ***The Lower Bound.*** Let $\varphi \in \mathscr{G}$ satisfying the assumptions of Theorem 16 and let us assume that around $\mathbf{1}$, Φ satisfies (24) and (25). Let $\nu > 0$. For $\delta \in (0, 1/\nu)$, we consider the sequence $S_\delta = (s_m)$, given by

$$s_m = \frac{1}{2} + \nu\delta + im\delta, \quad \text{where} \quad 1 \leq m \leq \Big(\frac{1}{\delta}\Big)^{1-\frac{1}{k_1}}.$$

We intend to apply part (b) of Lemma 21. We will require the construction of preimages of $S_\delta - \frac{1}{2}$ by Φ, and the inverse function theorem will provide the solution.

Lemma 25. *Let $\varphi \in \mathcal{G}$ satisfy the assumptions of Theorem 16. Then there exist $\nu_0, C_1, C_2 > 0$ such that for all $\nu \geq \nu_0$ and every $\delta \in (0, 1/\nu)$, there exists a finite sequence $Z_\delta = (Z(\alpha))$ in $\mathbb{D}^d$ with*

$$Z(\alpha) = [(1-\rho_1(\alpha))e^{i\theta_1(\alpha)}, \dots, (1-\rho_d(\alpha))e^{i\theta_d(\alpha)}]$$

such that

- *for any $\alpha \neq \beta$, we have $\sup_{j=1,\dots,d} |\theta_j(\alpha) - \theta_j(\beta)| \geq C_1\delta$,*
- *for any α and any $j = 1, \dots, d$, we have $C_2^{-1}\delta \leq \rho_j(\alpha) \leq C_2\delta$,*
- *$\Phi(Z_\delta) = S_\delta - \frac{1}{2}$ and, for any $1 \leq m \leq (1/\delta)^{1-\frac{1}{k_1}}$, the equation $\Phi(Z(\alpha)) = s_m - \frac{1}{2}$ has at least $\prod_{j=2}^{d} \left\lfloor (1/\delta)^{1-\frac{1}{k_j}} \right\rfloor$ solutions.*

Proof. We start as in the deduction of the upper bound in Theorem 16 from Theorem 22, writing

$$\operatorname{Re}\Phi(z) = a_1\rho_1 + \dots + a_d\rho_d + \ell_1(\theta)^{k_1} + \dots + \ell_d(\theta)^{k_d} + o\left(\sum_{j=1}^{d} \left(\rho_j + \ell_j^{k_j}(\theta)\right)\right),$$

$$\operatorname{Im}\Phi(z) = \tau + b_1\ell_1(\theta) + \dots + b_d\ell_d(\theta) + o\left(\sum_{j=1}^{d} |\ell_j(\theta)|\right),$$

for $z = ((1-\rho_1)e^{i\theta_1}, \dots, (1-\rho_d)e^{i\theta_d})$. To simplify the notations, we use the (linear) change of variables $u_j = \ell_j(\theta)$. We also set

$$\Lambda = \mathbb{N}^d \cap \prod_{j=1}^{d} \left[1, \left(\frac{1}{\delta}\right)^{1-\frac{1}{k_j}}\right]$$

and, for $\alpha \in \Lambda$ and $j = 2, \dots, d$ we let $\rho_j(\alpha) = \delta$ and $u_j(\alpha) = \alpha_j\delta$.

Setting $m = \alpha_1$, we want $Z(\alpha)$ such that $\operatorname{Re}\Phi(Z(\alpha)) = \nu\delta$ and $\operatorname{Im}\Phi(Z(\alpha)) = m\delta$. We are left to determine $\rho_1(\alpha)$ and $u_1(\alpha)$. We rewrite this system as

$$\begin{cases} f_\alpha(\rho_1, u_1)\rho_1 + g_\alpha(\rho_1, u_1)u_1^{k_1} & = \nu\delta + d_\alpha \\ h_\alpha(\rho_1, u_1)u_1 & = m\delta + e_\alpha \end{cases}, \tag{26}$$

where f_α, g_α and h_α are smooth functions depending only on $\alpha_2, \dots, \alpha_d$ and there exists a neighborhood $\mathcal{U} \ni (0,0)$ so that for every $(\rho, u) \in \mathcal{U}$,

$$|f_\alpha(\rho, u) - a_1| \ll \delta, \quad |g_\alpha(\rho, u) - 1| \ll \delta \quad \text{and} \quad |h_\alpha(\rho, u) - 1| \ll \delta.$$

Here, the open set $\mathcal{U}$ and the involved constants are uniform with respect to α, $\nu \geq 1$ and $\delta \in (0, 1/\nu)$. Moreover, the real numbers d_α and e_α satisfy

$$d_\alpha \ll \sum_{j=2}^{d} \delta + \sum_{j=2}^{d} \left(\left(\frac{1}{\delta}\right)^{1-\frac{1}{k_j}}\right)^{k_j} \ll \delta \quad \text{and} \quad e_\alpha \ll \delta \sum_{j=2}^{d} \left(\frac{1}{\delta}\right)^{1-\frac{1}{k_j}} \ll \delta^{\frac{1}{k_1}}.$$

We now apply the inverse function theorem to solve the system (26). Provided ν is large enough, we get a solution $(\rho_1(\alpha), u_1(\alpha))$ satisfying $\sup(\rho_1(\alpha), |u_1(\alpha)|) \ll \delta^{1/k_1}$. In this case, the involved constant depends on ν, but it is uniform with respect to α and δ.

Now, a look at the first equation of (26) shows that we in fact have the more precise inequality $\delta \ll \rho_1(\alpha) \ll \delta$, provided ν is sufficiently large, and this is independent of α and $\delta \in (0, 1)$. Looking now at the second equation of (26), if $\alpha \neq \beta \in \Lambda$ satisfy $\alpha_j = \beta_j$ for $j \geq 2$, so that $e_\alpha = e_\beta$ and $h_\alpha = h_\beta$, then $|u_1(\alpha) - u_1(\beta)| \gg \delta$.

Hence, we have obtained $\prod_{j=2}^{d} \lfloor (1/\delta)^{1-\frac{1}{k_j}} \rfloor$ solutions to the equation $\Phi(Z(\alpha)) = s_m$, and they satisfy the conclusions of Lemma 25 since the inequalities on $u_j(\alpha)$ are also valid for $\theta_j(\alpha)$ up to a constant depending only of Φ. □

Final part in the proof of the lower bound of Theorem 16. We apply Lemma 21 to S_δ and Z_δ given by the previous lemma, for

$$\delta = \left(\frac{1}{n}\right)^{\frac{k_1}{k_1-1}},$$

so that S_δ has cardinal number equal to n. Since

$$M_{\mathcal{H}^2}(S_\delta) \ll 1 \quad \text{and} \quad \left\| \mu_{Z,H^2(\mathbb{T}^d)} \right\|_{\mathscr{C},H^2(\mathbb{T}^d)} \ll 1$$

by Lemma 19 and Lemma 20, we are left to estimate the sum $N_\Phi(s_m; Z)\zeta(2\,\mathrm{Re}\, s_m)$ for any m. Using the fact that $\rho_j(\alpha) \gg \delta$ for any $j = 1, \dots, d$ and any α, we obtain

$$N_\Phi(s_m; Z)\zeta(2\mathrm{Re} s_m) \gg \left(\frac{1}{\delta}\right)^{\sum_{j=2}^{d}\left(1-\frac{1}{k_j}\right)} \cdot \delta^d \cdot \delta^{-1} \gg \delta^{\sum_{j=2}^{d} \frac{1}{k_j}} \gg \left(\frac{1}{n}\right)^{\left(\sum_{j=2}^{d} \frac{1}{k_j}\right) \times \frac{k_1}{k_1-1}},$$

and we are done. □

Remark. We may modify the proof of Theorem 16 so that we do not assume that there exists a unique $w \in \mathbb{T}^d$ such that $\mathrm{Re}\, \Phi(w) = 0$. Suppose that φ is boundary regular at any $w \in \mathbb{T}^d$ such that $\mathrm{Re}\, \Phi(w) = 0$. Define now the *compactness index* of φ as the real number

$$\eta_\varphi(s) = \inf\{\eta_{\varphi,w} : \mathrm{Re}\, \Phi(w) = 0\}.$$

It should be observed that this infimum is in fact a minimum. Indeed, our assumptions imply that the points $w \in \mathbb{T}^d$ such that $\mathrm{Re}\, \Phi(w) = 0$ are isolated. Theorem 16 remains true with this new definition of η_φ.

References

[Bayart 2002] F. Bayart, "Hardy spaces of Dirichlet series and their composition operators", *Monatsh. Math.* **136**:3 (2002), 203–236. MR Zbl

[Bayart 2003] F. Bayart, "Compact composition operators on a Hilbert space of Dirichlet series", *Illinois J. Math.* **47**:3 (2003), 725–743. MR Zbl

[Bayart and Brevig 2015] F. Bayart and O. F. Brevig, Xcas file for Lemma 11, 2015, available at http://bayart.perso.math.cnrs.fr/nonlin.xws.

[Berndtsson et al. 1987] B. Berndtsson, S.-Y. A. Chang, and K.-C. Lin, "Interpolating sequences in the polydisc", *Trans. Amer. Math. Soc.* **302**:1 (1987), 161–169. MR Zbl

[Bruce and Giblin 1992] J. W. Bruce and P. J. Giblin, *Curves and singularities: a geometrical introduction to singularity theory*, 2nd ed., Cambridge University Press, 1992. MR Zbl

[Chang 1979] S.-Y. A. Chang, "Carleson measure on the bi-disc", *Ann. of Math.* (2) **109**:3 (1979), 613–620. MR Zbl

[Finet and Queffélec 2004] C. Finet and H. Queffélec, "Numerical range of composition operators on a Hilbert space of Dirichlet series", *Linear Algebra Appl.* **377** (2004), 1–10. MR Zbl

[Finet et al. 2004] C. Finet, H. Queffélec, and A. Volberg, "Compactness of composition operators on a Hilbert space of Dirichlet series", *J. Funct. Anal.* **211**:2 (2004), 271–287. MR

[Gordon and Hedenmalm 1999] J. Gordon and H. Hedenmalm, "The composition operators on the space of Dirichlet series with square summable coefficients", *Michigan Math. J.* **46**:2 (1999), 313–329. MR Zbl

[Hardy and Wright 1979] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford University Press, New York, 1979. MR Zbl

[Queffélec and Seip 2015a] H. Queffélec and K. Seip, "Approximation numbers of composition operators on the H^2 space of Dirichlet series", *J. Funct. Anal.* **268**:6 (2015), 1612–1648. MR Zbl

[Queffélec and Seip 2015b] H. Queffélec and K. Seip, "Decay rates for approximation numbers of composition operators", *J. Anal. Math.* **125** (2015), 371–399. MR Zbl

Received November 17, 2015.

FRÉDÉRIC BAYART
LABORATOIRE DE MATHÉMATIQUES, UMR 6620, CNRS
CLERMONT UNIVERSITÉ, UNIVERSITÉ BLAISE PASCAL, CAMPUS DE CÉZEAUX
3, PLACE VASARELY, BP 10448
TSA 60026, CS 60026
63178 AUDIÈRE CEDEX
FRANCE

bayart@math.univ-bpclermont.fr

OLE FREDRIK BREVIG
DEPARTMENT OF MATHEMATICAL SCIENCES
NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY (NTNU)
7491 TRONDHEIM
NORWAY

ole.brevig@math.ntnu.no

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.121

A LOCAL RELATIVE TRACE FORMULA FOR PGL(2)

PATRICK DELORME AND PASCALE HARINCK

Following a scheme inspired by recent results of B. Feigon, who obtained what she called a local relative trace formula for PGL_2 and a local Kuznetsov trace formula for $U(2)$, we describe the spectral side of a local relative trace formula for $G := \mathrm{PGL}(2, E)$ relative to the symmetric subgroup $H := \mathrm{PGL}(2, F)$ where E/F is an unramified quadratic extension of local nonarchimedean fields of characteristic 0. The spectral side is given in terms of regularized normalized periods and normalized C-functions of Harish-Chandra. Using the geometric side of the local relative trace formula obtained in a more general setting by the authors and S. Souaifi, we deduce a local relative trace formula for G relative to H. We apply our result to invert some orbital integrals.

1. Introduction

Let E/F be an unramified quadratic extension of local nonarchimedean fields of characteristic 0. In this paper, we prove a local relative trace formula for $G := \mathrm{PGL}(2, E)$ relative to the symmetric subgroup $H := \mathrm{PGL}(2, F)$ following a scheme inspired by B. Feigon [2012].

As in [Arthur 1991], the way to establish a local relative trace formula is to describe two asymptotic expansions of a truncated kernel associated to the regular representation of $G \times G$ on $L^2(G)$, the first one in terms of weighted orbital integrals (called the geometric expansion), and the second one in terms of irreducible representations of G (called the spectral expansion). The truncated kernel we consider is defined as follows. The regular representation R of $G \times G$ on $L^2(G)$ is given by $(R(g_1, g_2)\psi)(x) = \psi(g_2^{-1} x g_1)$. For $f = f_1 \otimes f_2$, where f_1 and f_2 are two smooth compactly supported functions on G, the corresponding operator $R(f)$ is an integral operator on $L^2(G)$ with smooth kernel

$$K_f(x, y) = \int_G f_1(gy) f_2(xg)\, dg = \int_G f_1(x^{-1} gy) f_2(g)\, dg.$$

The first author was supported by a grant of Agence Nationale de la Recherche with reference ANR-13-BS01-0012 FERPLAY.
MSC2010: 11F72, 22E50.
Keywords: p-adic reductive groups, symmetric spaces, local relative trace formula, truncated kernel, regularized periods.

We define the truncated kernel $K^n(f)$ by

$$K^n(f) := \int_{H\times H} K_f(x, y)u(x, n)u(y, n)\, dx\, dy,$$

where the truncated function $u(\cdot, n)$ is the characteristic function of a large compact subset in H depending on a positive integer n as in [Arthur 1991] or [Delorme et al. 2015].

In the later reference, we studied such a truncated kernel in the more general setting where H is the group of F-points of a reductive algebraic group $\underline{H}$ defined and split over F and G is the group of F-points of the restriction of scalars $\underline{G} := \mathrm{Res}_{E/F}\, \underline{H}$ from E to F. We obtained an asymptotic geometric expansion of this truncated kernel in terms of weighted orbital integrals.

It is considerably more difficult to obtain a spectral asymptotic expansion of the truncated kernel and the main part of this paper is devoted to giving it for $\underline{H} = \mathrm{PGL}(2)$.

First, we express the kernel K_f in terms of normalized Eisenstein integrals using the Plancherel formula for G (see Section 3). Then the truncated kernel can be written as a finite linear combination, depending on unitary irreducible representations of G, of terms involving scalar product of truncated periods (see Corollary 4.2). The difficulty appears in the terms depending on principal series of G.

Let M and P be the images in G of the group of diagonal and upper triangular matrices of $\mathrm{GL}(2, E)$, respectively, and let $\bar{P}$ be the parabolic subgroup opposite to P. As M is isomorphic to $E^\times$, we identify characters on M and on $E^\times$. The group of unramified characters of M is isomorphic to $\mathbb{C}^*$ by a map $z \to \chi_z$. Let δ be a unitary character of $E^\times$, which is trivial on a fixed uniformizer of $F^\times$. For $z \in \mathbb{C}^*$, we set $\delta_z := \delta \otimes \chi_z$. We denote by $(i_P^G \delta_z, i_P^G \mathbb{C}_{\delta_z})$ the normalized induced representation and by $(i_P^G \check{\delta}_z, i_P^G \check{\mathbb{C}}_{\delta_z})$ its contragredient. Then, the normalized truncated period is defined by

$$P^n_{\delta_z}(S) := \int_H E^0(P, \delta_z, S)(h)u(h, n)\, dh, \quad S \in i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z},$$

where $E^0(P, \delta_z, \cdot)$ is the normalized Eisenstein integral associated to $i_P^G \delta_z$ (see (3-6)). The contribution of $i_P^G \delta_z$ in $K^n(f)$ is a finite linear combination of integrals

$$I^n_\delta(S, S') := \int_{\mathcal{O}} P^n_{\delta_z}(S)\overline{P^n_{\delta_z}(S')}\, \frac{dz}{z}, \quad S, S' \in i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z},$$

where $\mathcal{O}$ is the torus of complex numbers of modulus equal to 1.

To establish the asymptotic expansion of this integral, we recall the notion of normalized regularized period introduced by Feigon (see Section 4). This period, denoted by

$$P_{\delta_z}(S) := \int_H^* E^0(P, \delta_z, S)(h)\, dh$$

is meromorphic in a neighborhood $\mathcal{V}$ of $\mathcal{O}$ with at most a simple pole at $z = 1$ and defines an $H \times H$ invariant linear form on $i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}$. Moreover, the difference $P_{\delta_z}(S) - P_{\delta_z}^n(S)$ is a rational function in z on $\mathcal{V}$ with at most a simple pole at $z = 1$ which depends on the normalized C-functions of Harish-Chandra. As normalized Eisenstein integrals and normalized C-functions are holomorphic in a neighborhood of $\mathcal{O}$, we can deduce an asymptotic behavior of the integrals $I_\delta^n(S, S')$ in terms of normalized regularized periods and normalized C-functions (see Proposition 7.1).

Our first result (see Theorem 7.3) asserts that $K^n(f)$ is asymptotic to a polynomial function in n of degree 1 whose coefficients are described in terms of generalized matrix coefficients $m_{\xi,\xi'}$ associated to unitary irreducible representations (π, V_π) of G where ξ and ξ' are linear forms on V_π. When (π, V_π) is a normalized induced representation, these linear forms are defined from the regularized normalized periods, its residues, and the normalized C-functions of Harish-Chandra.

We make precise the geometric asymptotic expansion of $K^n(f)$ obtained in [Delorme et al. 2015] for $\underline{H} := \mathrm{PGL}(2)$. Therefore, comparing the two asymptotic expansions of $K^n(f)$, we deduce our relative local trace formula and a relation between orbital integrals on elliptic regular points in $H\backslash G$ and some generalized matrix coefficients of induced representations (Theorem 8.1).

As corollaries of these results, we give an inversion formula for orbital integrals on regular elliptic points of $H\backslash G$ and for orbital integrals of a matrix coefficient associated to a cuspidal representation of G.

2. Notation

Let F be a nonarchimedean local field of characteristic 0 and odd residual characteristic q. Let E be an unramified quadratic extension of F. Let $\mathcal{O}_F$ and $\mathcal{O}_E$ denote the rings of integers in F and E. We fix a uniformizer ω in the maximal ideal of $\mathcal{O}_F$. Thus ω is also a uniformizer of E. We denote by $v(\cdot)$ the valuation of F, extended to E. Let $|\cdot|_F$ and $|\cdot|_E$ denote the normalized valuations on F and E. Thus for $a \in F^\times$, one has $|a|_F = |a|_E^2$.

Let $N_{E/F}$ be the norm map from $E^\times$ to $F^\times$. We denote by E^1 the set of elements in $E^\times$ whose norm is equal to 1.

Let $\underline{H} := \mathrm{PGL}(2)$ defined over F and let $\underline{G} := \mathrm{Res}_{E/F}(\underline{H} \times_F E)$ be the restriction of scalars of $\underline{H}$ from E to F. We set $H := \underline{H}(F) = \mathrm{PGL}(2, F)$ and $G := \underline{G}(F) = \mathrm{PGL}(2, E)$. Let $K := \underline{G}(\mathcal{O}_F) = \mathrm{PGL}(2, \mathcal{O}_E)$.

We denote by $C^\infty(G)$ the space of smooth functions on G and by $C_c^\infty(G)$ the subspace of compactly supported functions in $C^\infty(G)$. If V is a vector space of valued functions on G which is invariant by right and left translation, we will denote by ρ and λ, respectively, the right and left regular representation of G in V.

If V is a vector space, V' will denote its dual. If V is real, $V_{\mathbb{C}}$ will denote its complexification.

Let p be the canonical projection of $\mathrm{GL}(2, E)$ onto G. We denote by M and N the image by p of the subgroups of diagonal matrices and upper triangular unipotent matrices of $\mathrm{GL}(2, E)$, respectively. We set $P := MN$ and we denote by $\bar{P}$ the parabolic subgroup opposite to P. Let δ_P be the modular function of P. We denote by 1 and w the representatives in K of the Weyl group W^G of M in G.

For $J = K, M$ or P, we set $J_H := J \cap H$.

For a, b in $E^\times$, we denote by $\mathrm{diag}_G(a, b)$ the image by p of the diagonal matrix $\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right) \in \mathrm{GL}(2, E)$. The natural map $(a, b) \mapsto \mathrm{diag}_G(a, b)$ induces an isomorphism from $E^\times \times E^\times / \mathrm{diag}(E^\times) \simeq E^\times$ to M where $\mathrm{diag}(E^\times)$ is the diagonal of $E^\times \times E^\times$. *Hence, each character χ of $E^\times$ defines a character of M given by*

$$\mathrm{diag}_G(a, b) \mapsto \chi(ab^{-1}), \tag{2-1}$$

which we will denote by the same letter. We define the map $h_M : M \to \mathbb{R}$ by

$$q^{-h_M(m)} = |ab^{-1}|_E, \quad \text{for } m = \mathrm{diag}_G(a, b). \tag{2-2}$$

We define similarly h_{M_H} on M_H by $q^{-h_{M_H}(\mathrm{diag}_G(a,b))} = |ab^{-1}|_F$ for $a, b \in F^\times$. Then for $m \in M_H$, one has $\delta_P(m) = \delta_{P_H}(m)^2 = q^{-2h_{M_H}(m)}$.

We normalize the Haar measure dx on F so that $\mathrm{vol}(\mathcal{O}_F) = 1$. We define the measure $d^\times x$ on $F^\times$ by

$$d^\times x = \frac{1}{1 - q^{-1}} \frac{1}{|x|_F} dx.$$

Thus, we have $\mathrm{vol}(\mathcal{O}_F^\times) = 1$. We let M and M_H have the measure induced by $d^\times x$. We normalize the Haar measure on K so that $\mathrm{vol}(K) = 1$. Let dn be the Haar measure on N such that

$$\int_N \delta_{\bar{P}}(m_{\bar{P}}(n))\, dn = 1.$$

Let dg be the Haar measure on G such that

$$\int_G f(g) dg = \int_M \int_N \int_K f(mnk)\, dk\, dn\, dm.$$

We define dh on H similarly.

The Cartan decomposition of H is given by

$$H = K_H M_H^+ K_H, \quad \text{where } M_H^+ := \{\mathrm{diag}_G(a, b); a, b \in F^\times, |ab^{-1}|_F \le 1\}, \tag{2-3}$$

and for any integrable function f on H, we have the standard integration formula

$$\int_H f(x)\,dx = \int_{K_H}\int_{K_H}\int_{M_H} D_{P_H}(m) f(k_1 m k_2)\,dm\,dk_2\,dk_1, \tag{2-4}$$

where

$$D_{P_H}(m) = \begin{cases} \delta_{P_H}(m)^{-1}(1+q^{-1}) & \text{if } m \in M_H^+, \\ 0 & \text{otherwise.} \end{cases}$$

For $h \in H$, we denote by $\mathcal{M}(h)$ an element of M_H^+ such that $h \in K_H \mathcal{M}(h) K_H$. The element $h_{M_H}(\mathcal{M}(h))$ is independent of this choice. We thank E. Lapid, who suggested the proof of the following lemma.

Lemma 2.1. *Let Ω be a compact subset of H. There is an $N_0 > 0$ satisfying the following property: for any $h \in \Omega$, there exists $X_h \in \mathbb{R}$ such that, for all $m \in M_H^+$ satisfying $h_{M_H}(m) \geq N_0$, one has*

$$h_{M_H}(\mathcal{M}(mh)) = h_{M_H}(m) + X_h.$$

Proof. For a matrix $x = (x_{i,j})_{i,j}$ of $\mathrm{GL}(2, F)$, we set

$$F(x) := \log \max_{i,j} \left(\frac{|x_{i,j}|_F^2}{|\det x|_F} \right).$$

The function F is clearly invariant under the action of the center of $\mathrm{GL}(2, F)$, hence it defines a function on H which we denote by the same letter.

Since $|\cdot|_F$ is ultrametric, for $k \in K_H$ and $h \in H$, we have $F(kh) \leq F(h)$, hence $F(k^{-1}kh) \leq F(kh)$. Using the same argument on the right, we deduce that F is right and left invariant by K_H.

If $m = \mathrm{diag}_G(\omega^{n_1}, \omega^{n_2})$ with $n_1 - n_2 \geq 0$ then

$$F(m) = \log \max\left(\frac{q^{-2n_1}}{q^{-n_1-n_2}}, \frac{q^{-2n_2}}{q^{-n_1-n_2}} \right) = (n_1 - n_2)\log q = h_{M_H}(m)\log q.$$

Thus, we deduce that $F(h) = h_{M_H}(\mathcal{M}(h))\log q$, for $h \in H$.

If $h = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $m = \mathrm{diag}_G(\omega^{n_1}, \omega^{n_2})$, then

$$F(mh) = \log\max(|a|_F^2 q^{n_2-n_1}, |b|_F^2 q^{n_2-n_1}, |c|_F^2 q^{n_1-n_2}, |d|_F^2 q^{n_1-n_2}) - \log|ad-bc|_F.$$

Therefore, we can choose $N_0 > 0$ such that, for any $h \in \Omega$ and $m \in M_H^+$ with $h_M(m) > N_0$, we have

$$\begin{aligned} F(mh) &= \log\max(|c|_F^2 q^{n_1-n_2}, |d|_F^2 q^{n_1-n_2}) - \log|ad-bc|_F \\ &= (n_1 - n_2)\log q + \log\max(|c|_F^2, |d|_F^2) - \log|ad-bc|_F. \end{aligned}$$

Hence, we obtain the lemma. $\square$

3. Normalized Eisenstein integrals and Plancherel formula

We denote by $\widehat{M}_2$ the set of unitary characters of $E^\times$ which are trivial on ω.

Let $X(M)$ be the complex torus of unramified characters of M and $X(M)_u$ be the compact subtorus of unitary unramified characters of M. For $z \in \mathbb{C}^*$, we denote by χ_z the unramified character of $E^\times$ defined by $\chi_z(\omega) = z$. By definition of h_M, we have $\chi_z(m) = z^{h_M(m)/2}$. Each element of $X(M)$ is of the form χ_z for some $z \in \mathbb{C}^*$ and $X(M)_u$ identifies with the group $\mathcal{O}$ of complex numbers of modulus equal to 1. For $\delta \in \widehat{M}_2$ and $z \in \mathbb{C}^*$, we set $\delta_z := \delta \otimes \chi_z$. We will denote by $\mathbb{C}_{\delta_z}$ the space of δ_z.

Let $Q = MU$ be equal to P or to $\bar{P}$. Let $\delta \in \widehat{M}_2$ and $z \in \mathbb{C}^*$. We denote by $i_Q^G \delta_z$ the right representation of G in the space $i_Q^G \mathbb{C}_{\delta_z}$ of maps v from G to $\mathbb{C}$, right invariant by a compact open subgroup of G such that $v(mug) = \delta_Q(m)^{1/2}\delta_z(m)v(g)$ for all $m \in M$, $u \in U$ and $g \in G$.

We denote by $(\bar{i}_Q^G \delta_z, i_{K\cap Q}^K \mathbb{C})$ the compact realization of $(i_Q^G \delta_z, i_Q^G \mathbb{C}_{\delta_z})$ obtained by restriction of functions. If $v \in i_{Q\cap K}^K \mathbb{C}$, we denote by v_z the element of $i_Q^G \mathbb{C}_{\delta_z}$ whose restriction to K is equal to v.

We define a scalar product on $i_{Q\cap K}^K \mathbb{C}$ by

$$(v, v') = \int_K v(k)\overline{v'(k)}\, dk, \quad v, v' \in i_{Q\cap K}^K \mathbb{C}. \tag{3-1}$$

If $z \in \mathcal{O}$ (hence δ_z is unitary), the representation $\bar{i}_Q^G(\delta_z)$ is unitary. Therefore, by "transport de structure", $i_Q^G(\delta_z)$ is also unitary.

Let $(\check{\delta}_z, \check{\mathbb{C}}_{\delta_z})$ be the contragredient representation of $(\delta_z, \mathbb{C}_{\delta_z})$. We can and will identify $(i_Q^G \check{\delta}_z, i_Q^G \check{\mathbb{C}}_{\delta_z})$ with the contragredient representation of $(i_Q^G \delta_z, i_Q^G \mathbb{C}_{\delta_z})$ and $i_Q^G \mathbb{C}_{\delta_z} \otimes i_Q^G \check{\mathbb{C}}_{\delta_z}$ with a subspace of $\mathrm{End}_G(i_Q^G \mathbb{C}_{\delta_z})$ [Waldspurger 2003, I.3].

Using the isomorphism between $i_Q^G \mathbb{C}_{\delta_z}$ and $i_{Q\cap K}^K \mathbb{C}$, we can define the notion of rational or polynomial map from $X(M)$ to a space depending on $i_Q^G \mathbb{C}_{\delta_z}$ as in [ibid., IV.1 and VI.1].

We denote by $A(\bar{Q}, Q, \delta_z) : i_Q^G \mathbb{C}_{\delta_z} \to i_{\bar{Q}}^G \mathbb{C}_{\delta_z}$ the standard intertwining operator. By [ibid., IV.1 and Proposition IV.2.2], the map $z \in \mathbb{C}^* \mapsto A(\bar{Q}, Q, \delta_z) \in \mathrm{Hom}_G(i_Q^G \mathbb{C}_{\delta_z}, i_{\bar{Q}}^G \mathbb{C}_{\delta_z})$ is a rational function on $\mathbb{C}^*$. Moreover, there exists a rational complex valued function $j(\delta_z)$ depending only on M such that $A(Q, \bar{Q}, \delta_z) \circ A(\bar{Q}, Q, \delta_z)$ is the dilation of scale $j(\delta_z)$. We set

$$\mu(\delta_z) := j(\delta_z)^{-1}. \tag{3-2}$$

By [ibid., Lemme V.2.1], the map $z \mapsto \mu(\delta_z)$ is rational on $\mathbb{C}^*$ and regular on $\mathcal{O}$.

The Eisenstein integral $E(Q, \delta_z)$ is the map from $i_Q^G \mathbb{C}_{\delta_z} \otimes i_Q^G \check{\mathbb{C}}_{\delta_z}$ to $C^\infty(G)$ defined by

$$E(Q, \delta_z, v \otimes \check{v})(g) = \langle (i_Q^G \delta_z)(g)v, \check{v}\rangle, \quad v \in i_Q^G \mathbb{C}_{\delta_z}, \check{v} \in i_Q^G \check{\mathbb{C}}_{\delta_z}. \tag{3-3}$$

If $\psi \in i_Q^G \mathbb{C}_{\delta_z} \otimes i_Q^G \check{\mathbb{C}}_{\delta_z}$ is identified with an endomorphism of $i_Q^G \mathbb{C}_{\delta_z}$, we have

$$E(Q, \delta_z, \psi)(g) = \mathrm{tr}(i_Q^G \delta_z(g)\psi). \tag{3-4}$$

We introduce the operator $C_{P,P}(1, \delta_z) := \mathrm{Id} \otimes A(\bar{P}, P, \check{\delta}_z)$ from $i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}$ to $i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}$. By [Waldspurger 2003, Lemme V.2.2]

(3-5) the operator $\mu(\delta_z)^{1/2} C_{P,P}(1, \delta_z)$ is unitary and regular on $\mathcal{O}$.

We define the normalized Eisenstein integral $E^0(P, \delta_z) : i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z} \to C^\infty(G)$ by

$$E^0(P, \delta_z, \Psi) = E(P, \delta_z, C_{P|P}(1, \delta_z)^{-1}\Psi). \tag{3-6}$$

By [Silberger 1979, §5.3.5]

(3-7) $E^0(P, \delta_z, \Psi)$ is regular on $\mathcal{O}$.

For $f \in C_c^\infty(G)$, we denote by $\check{f}$ the function defined by $\check{f}(g) := f(g^{-1})$. Then, the operator $i_P^G \delta_z(\check{f})$ belongs to $i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z} \subset \mathrm{End}_G(i_P^G \mathbb{C}_{\delta_z})$. We define the Fourier transform $\mathcal{F}(P, \delta_z, f) \in i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}$ of f by

$$\mathcal{F}(P, \delta_z, f) = i_P^G \delta_z(\check{f}).$$

The G-invariant scalar product on $i_P^G \mathbb{C}_{\delta_z}$ defined in (3-1) induces a G-invariant scalar product on $i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}$ given by

$$(v_1 \otimes \check{v}_1, v_2 \otimes \check{v}_2) = (v_1, v_2)(\check{v}_1, \check{v}_2).$$

Notice that by the inclusion $i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z} \subset \mathrm{End}(i_P^G \mathbb{C}_{\delta_z})$, this scalar product coincides with the Hilbert–Schmidt scalar product on the space of Hilbert–Schmidt operators on $i_P^G \mathbb{C}_{\delta_z}$ defined by

$$(S, S') = \mathrm{tr}(SS'^*), \tag{3-8}$$

where $\mathrm{tr}(SS'^*) = \sum_{o.n.b.} \langle SS'^* u_i, u_i \rangle$ and this sum converges absolutely and does not depend on the basis. Then, the Fourier transform is the unique element of $i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}$ such that

$$(E(P, \delta_z, \Psi), f)_G = (\Psi, \mathcal{F}(P, \delta_z, f)). \tag{3-9}$$

Moreover, we have [Waldspurger 2003, Lemme VII.1.1]

$$E(P, \delta_z, \mathcal{F}(P, \delta_z, f))(g) = \mathrm{tr}[(i_P^G \delta_z)(\lambda(g)\check{f})]. \tag{3-10}$$

We define the normalized Fourier transform $\mathcal{F}^0(P, \delta_z, f)$ of $f \in C_c^\infty(G)$ as the unique element of $i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}$ such that

$$(\Psi, \mathcal{F}^0(P, \delta_z, f)) = (E^0(P, \delta_z, \Psi), f)_G, \quad \Psi \in i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}.$$

It follows easily from (3-9) and (3-5) that

$$\mathcal{F}^0(P, \delta_z, f) = \mu(\delta_z) C_{P|P}(1, \delta_z) \mathcal{F}(P, \delta_z, f),$$

thus we deduce that

$$E^0(P, \delta_z, \mathcal{F}^0(P, \delta_z, f)) = \mu(\delta_z) E(P, \delta_z, \mathcal{F}(P, \delta_z, f)). \tag{3-11}$$

Therefore, we can describe the spectral decomposition of the regular representation $R := \rho \otimes \lambda$ of $G \times G$ on $L^2(G)$ of [Waldspurger 2003, Théorème VIII.1.1] in terms of normalized Eisenstein integrals as follows. Let $\mathcal{E}_2(G)$ be the set of classes of irreducible admissible representations of G whose matrix coefficients are square-integrable. We will denote by $d(\tau)$ the formal degree of $\tau \in \mathcal{E}_2(G)$. Then we have (3-12)

$$f(g) = \sum_{\tau \in \mathcal{E}_2(G)} d(\tau) \operatorname{tr}(\tau(\lambda(g) \check{f})) + \frac{1}{4i\pi} \sum_{\delta \in \widehat{M}_2} \int_{\mathcal{O}} E^0(P, \delta_z, \mathcal{F}^0(P, \delta_z, f))(g) \frac{dz}{z}.$$

4. The truncated kernel

Let $f \in C_c^\infty(G \times G)$ be of the form $f(y_1, y_2) = f_1(y_1) f_2(y_2)$ with $f_j \in C_c^\infty(G)$. Then the operator $R(f)$ (where $R := \rho \otimes \lambda$) is an integral operator with smooth kernel

$$K_f(x, y) = \int_G f_1(gy) f_2(xg) \, dg = \int_G f_1(x^{-1} g y) f_2(g) \, dg.$$

Notice that the kernel studied in [Arthur 1991; Feigon 2012; Delorme et al. 2015] corresponds to the kernel of the representation $\lambda \times \rho$ which coincides with $K_{f_2 \otimes f_1}(x, y) = K_{f_1 \otimes f_2}(x^{-1}, y^{-1})$.

The aim of this part is to give a spectral expansion of the truncated kernel obtained by integrating K_f against a truncated function on $H \times H$ as in [Arthur 1991].

Lemma 4.1. *For $(\tau, V_\tau) \in \mathcal{E}_2(G)$, we fix an orthonormal basis $\mathcal{B}_\tau$ of the space of Hilbert–Schmidt operators on V_τ. For $\delta \in \widehat{M}_2$ and $z \in \mathcal{O}$, we fix an orthonormal basis $\mathcal{B}_{\bar{P},P}(\mathbb{C})$ of $i_{P \cap K}^K \mathbb{C} \otimes i_{\bar{P} \cap K}^K \check{\mathbb{C}}$. Using the isomorphism $S \mapsto S_z$ between $i_{P \cap K}^K \mathbb{C} \otimes i_{\bar{P} \cap K}^K \check{\mathbb{C}}$ and $i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}$, we have*

$$\begin{aligned} K_f(x, y) = &\sum_{\tau \in \mathcal{E}_2(G)} \sum_{S \in \mathcal{B}_\tau} d(\tau) \operatorname{tr}(\tau(x) \tau(f_1) S \tau(\check{f}_2)) \overline{\operatorname{tr}(\tau(y) S)} \\ &+ \frac{1}{4i\pi} \sum_{\delta \in \widehat{M}_2} \sum_{S \in \mathcal{B}_{\bar{P},P}(\mathbb{C})} \int_{\mathcal{O}} E^0(P, \delta_z, \Pi_{\delta_z}(f) S_z)(x) \overline{E^0(P, \delta_z, S_z)(y)} \frac{dz}{z}, \end{aligned}$$

where $\Pi_{\delta_z}(f) S_z := (i_P^G \delta_z \otimes i_{\bar{P}}^G \check{\delta}_z)(f) S_z = (i_P \delta_z)(f_1) S_z (i_{\bar{P}} \delta_z)(\check{f}_2)$ and the sums over S are all finite.

Proof. For $x \in G$, we set

$$h(v) := \int_G f_1(uvx) f_2(xu)\, du,$$

so that

$$K_f(x, y) = [\rho(yx^{-1})h](e). \tag{4-1}$$

If π is a representation of G, one has

$$\begin{aligned}\pi(\rho(yx^{-1})h) &= \int_{G\times G} f_1(ugy) f_2(xu)\pi(g)\, du\, dg\\ &= \int_{G\times G} f_1(u_1) f_2(xu)\pi(u^{-1}u_1 y^{-1})\, du\, du_1\\ &= \int_{G\times G} f_1(u_1) f_2(u_2)\pi(u_2^{-1}xu_1 y^{-1})\, du_1\, du_2\\ &= \pi(\check{f}_2)\pi(x)\pi(f_1)\pi(y^{-1}).\end{aligned}$$

Therefore, using the Hilbert–Schmidt scalar product (3-8), one obtains for $\tau \in \mathcal{E}_2(G)$,

$$\begin{aligned}\operatorname{tr} \tau(\rho(yx^{-1})h) &= \operatorname{tr} \tau(\check{f}_2)\tau(x)\tau(f_1)\tau(y)^* = (\tau(\check{f}_2)\tau(x)\tau(f_1), \tau(y))\\ &= \sum_{S\in\mathcal{B}_\tau} (\tau(\check{f}_2)\tau(x)\tau(f_1), S^*)\overline{(\tau(y), S^*)}\\ &= \sum_{S\in\mathcal{B}_\tau} \operatorname{tr}(\tau(x)\tau(f_1)S\tau(\check{f}_2))\overline{\operatorname{tr}(\tau(y)S)},\end{aligned} \tag{4-2}$$

where the sum over S in $\mathcal{B}_\tau$ is finite.

We consider now $\pi := i_P^G\delta_z$ with $\delta \in \widehat{M}_2$ and $z \in \mathcal{O}$. By (3-10) and (3-11), we have

$$E^0(P, \delta_z, \mathcal{F}^0(P, \delta_z, [\rho(yx^{-1})h]^{\vee})(e) = \mu(\delta_z) \operatorname{tr} \pi(\rho(yx^{-1})h). \tag{4-3}$$

Let $\mathcal{B}_{P,P}(\mathbb{C}_{\delta_z})$ be an orthonormal basis of $i_P^G\mathbb{C}_{\delta_z} \otimes i_P^G\check{\mathbb{C}}_{\delta_z}$. Since $f_1, f_2 \in C_c^\infty(G)$, the operators $\pi(f_1)$ and $\pi(\check{f}_2)$ are of finite rank. Therefore, we deduce as above that

$$\begin{aligned}\operatorname{tr} \pi(\rho(yx^{-1})h) &= \operatorname{tr}(\pi(\check{f}_2)\pi(x)\pi(f_1)\pi(y)^{-1})\\ &= \sum_{S\in\mathcal{B}_{P,P}(\mathbb{C}_{\delta_z})} \operatorname{tr}(\pi(x)\pi(f_1)S\pi(\check{f}_2))\overline{\operatorname{tr}(\pi(y)S)},\end{aligned}$$

where the sum over S in $\mathcal{B}_{P,P}(\mathbb{C}_{\delta_z})$ is finite.

In what follows, the sums over elements of an orthonormal basis will be always finite. Hence, by (3-4), we deduce that

$$\operatorname{tr} \pi(\rho(yx^{-1})h) = \sum_{S\in\mathcal{B}_{P,P}(\mathbb{C}_{\delta_z})} E(P, \delta_z, \pi(f_1)S\pi(\check{f}_2))(x)\overline{E(P, \delta_z, S)(y)}. \tag{4-4}$$

Recall that we fixed an orthonormal basis $\mathcal{B}_{\bar{P},P}(\mathbb{C})$ of the space $i_{P\cap K}^{K}\mathbb{C}\otimes i_{\bar{P}\cap K}^{K}\check{\mathbb{C}}$ which is isomorphic to $i_P^G\mathbb{C}_{\delta_z}\otimes i_{\bar{P}}^G\check{\mathbb{C}}_{\delta_z}$ by the map $S\mapsto S_z$. By (3-5), the family $\widetilde{S}(\delta_z):=\mu(\delta_z)^{-1/2}C_{P,P}(1,\delta_z)^{-1}S_z$ for $S\in\mathcal{B}_{\bar{P},P}(\mathbb{C})$ is an orthonormal basis of $i_P^G\mathbb{C}_{\delta_z}\otimes i_P^G\check{\mathbb{C}}_{\delta_z}$.

Moreover, using the inclusion $i_P^G\mathbb{C}_{\delta_z}\otimes i_{\bar{P}}^G\check{\mathbb{C}}_{\delta_z}\subset \operatorname{Hom}_G(i_{\bar{P}}^G\mathbb{C}_{\delta_z}, i_P^G\mathbb{C}_{\delta_z})$, and the adjunction property of the intertwining operator [Waldspurger 2003, IV.1.(11)], we have $C_{P,P}(1,\delta_z)^{-1}S=S\circ A(P,\bar{P},\delta_z)^{-1}$, for all $S\in i_P^G\mathbb{C}_{\delta_z}\otimes i_{\bar{P}}^G\check{\mathbb{C}}_{\delta_z}$. Since $A(P,\bar{P},\delta_z)^{-1}\circ i_P^G(\delta_z)=i_{\bar{P}}^G(\delta_z)\circ A(P,\bar{P},\delta_z)^{-1}$, writing (4-4) for the basis $\widetilde{S}(\delta_z)$, we obtain

$$\begin{aligned}
&\operatorname{tr}\pi(\rho(yx^{-1})h)\\
&\quad=\mu(\delta_z)^{-1}\sum_{S\in\mathcal{B}_{\bar{P},P}(\mathbb{C})}E(P,\delta_z,\pi(f_1)C_{P,P}(1,\delta_z)^{-1}(S_z)\pi(\check{f}_2))(x)\\
&\qquad\qquad\qquad\qquad\times\overline{E(P,\delta_z,C_{P,P}(1,\delta_z)^{-1}S_z)(y)}\\
&\quad=\mu(\delta_z)^{-1}\sum_{S\in\mathcal{B}_{\bar{P},P}(\mathbb{C})}E(P,\delta_z,C_{P,P}(1,\delta_z)^{-1}[(i_P^G\delta_z)(f_1)S_z(i_{\bar{P}}^G\delta_z)(\check{f}_2)])(x)\\
&\qquad\qquad\qquad\qquad\times\overline{E(P,\delta_z,C_{P,P}(1,\delta_z)^{-1}S_z)(y)}\\
&\quad=\mu(\delta_z)^{-1}\sum_{S\in\mathcal{B}_{\bar{P},P}(\mathbb{C})}E^0(P,\delta_z,(i_P^G\delta_z)(f_1)S_z(i_{\bar{P}}^G\delta_z)(\check{f}_2))(x)\overline{E^0(P,\delta_z,S_z)(y)}.
\end{aligned}$$

We set $\Pi_{\delta_z}:=i_P^G\delta_z\otimes i_{\bar{P}}^G\check{\delta}_z$. Then we have

$$\Pi_{\delta_z}(f)S_z=(i_P^G\delta_z)(f_1)S_z(i_{\bar{P}}^G\delta_z)(\check{f}_2). \tag{4-5}$$

By (4-3), we obtain

$$\begin{aligned}
&E^0(P,\delta_z,\mathcal{F}^0(P,\delta_z,[\rho(yx^{-1})h]\check{\ }))(e)\\
&\qquad\qquad=\sum_{S\in\mathcal{B}_{\bar{P},P}(\mathbb{C})}E^0(P,\delta_z,\Pi_{\delta_z}(f)S_z)(x)\overline{E^0(P,\delta_z,S_z)(y)}.
\end{aligned}$$

The lemma follows from (3-12), (4-1), (4-2) and the above result. □

To integrate the kernel K_f on $H\times H$, we introduce truncation as in [Arthur 1991]. Let n be a positive integer. Let $u(\cdot,n)$ be the truncated function defined on H by

$$u(h,n)=\begin{cases}1 & \text{if } h=k_1mk_2 \text{ with } k_1,k_2\in K_H, m\in H \text{ such that } 0\le |h_{M_H}(m)|\le n,\\ 0 & \text{otherwise.}\end{cases}$$

We define the truncated kernel by

$$K^n(f):=\int_{H\times H}K_f(x,y)u(x,n)u(y,n)\,dx\,dy. \tag{4-6}$$

Since $K_f(x^{-1}, y^{-1})$ coincides with the kernel studied in [Delorme et al. 2015, 2.2] and $u(x,n) = u(x^{-1}, n)$, this definition of the truncated kernel coincides with the one in that reference.

We define truncated periods by

$$P_\tau^n(S) := \int_H \operatorname{tr}(\tau(y)S)u(y,n)\,dy, \quad (\tau, V_\tau) \in \mathcal{E}_2(G),\ S \in \operatorname{End}_{\text{fin.rk}}(V_\tau), \tag{4-7}$$

where $\operatorname{End}_{\text{fin.rk}}(V_\tau)$ is the space of finite rank operators in $\operatorname{End}(V_\tau)$, and

$$P_{\delta_z}^n(S) := \int_H E^0(P, \delta_z, S_z)(y)u(y,n)\,dy, \qquad \delta \in \widehat{M}_2,\ z \in \mathcal{O},\ S \in i_{P\cap K}^K \mathbb{C} \otimes i_{\bar{P}\cap K}^K \check{\mathbb{C}}. \tag{4-8}$$

Corollary 4.2. *With the notation of Lemma 4.1, one has*

$$K^n(f) = \sum_{\tau\in\mathcal{E}_2(G)} \sum_{S\in\mathcal{B}_\tau} d(\tau) P_\tau^n(\tau\otimes\check{\tau}(f)S)\overline{P_\tau^n(S)} + \frac{1}{4i\pi} \sum_{\delta\in\widehat{M}_2} \sum_{S\in\mathcal{B}_{\bar{P},P}(E)} \int_{\mathcal{O}} P_{\delta_z}^n(\overline{\Pi}_{\delta_z}(f)S)\overline{P_{\delta_z}^n(S)}\,\frac{dz}{z},$$

where the sums over S are all finite and $\overline{\Pi}_{\delta_z} := \bar{i}_P^G \delta_z \otimes \bar{i}_{\bar{P}}^G \check{\delta}_z$.

Proof. For $\tau \in \mathcal{E}_2(G)$ and $S \in \mathcal{B}_\tau$, one has

$$\tau(f_1)S\tau(\check{f}_2) = \tau\otimes\check{\tau}(f)S.$$

Therefore, since the functions we integrate are compactly supported, the assertion follows from Lemma 4.1. □

5. Regularized normalized periods

To determine the asymptotic expansion of the truncated kernel, we recall the notion of regularized period introduced in [Feigon 2012]. It is defined by meromorphic continuation.

Let $z_0 \in \mathbb{C}^*$. Then, for $z \in \mathbb{C}^*$ such that $|zz_0| < 1$, the integral

$$\int_{M_H^+} \chi_{z_0}(m)\chi_z(m)(1-u(m,n_0))\,dm = \sum_{n>n_0} (zz_0)^n = \frac{(zz_0)^{n_0+1}}{1-zz_0}$$

is well defined and has a meromorphic continuation at $z = 1$. Moreover this meromorphic continuation is holomorphic on $\mathcal{V} - \{1\}$ with a simple pole at $z_0 = 1$.

Let $\delta \in \widehat{M}_2$. We consider now an holomorphic function $z \mapsto \varphi_z \in C^\infty(G)$ defined in a neighborhood $\mathcal{V}$ of $\mathcal{O}$ in $\mathbb{C}^*$ such that

(5-1) there exist an integer $n_0 > 0$ and holomorphic functions $\mathcal{V} \to C^\infty(K_H \times K_H)$, $z \mapsto \phi_z^i$, $i = 1, 2$, such that

$$\delta_P(m)^{-1/2}\varphi_z(k_1 m k_2) = \delta_z(m)\phi_z^1(k_1, k_2) + \delta_{z^{-1}}(m)\phi_z^2(k_1, k_2),$$

for $k_1, k_2 \in K_H$, and $m \in M_H^+$ satisfying $h_{M_H}(m) > n_0$,

Recall that $\mathcal{M}(h)$ for $h \in H$ is an element in M_H^+ such that $h \in K_H \mathcal{M}(h) K_H$. By the integral formula (2-4), we deduce that for $|z| < \min(|z_0|, |z_0|^{-1})$, the integral

$$\begin{aligned}\int_H & \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))(1 - u(h, n_0))\, dh \\ &= (1 + q^{-1})\left(\int_{K_H \times K_H} \phi_{z_0}^1(k_1, k_2)\, dk_1\, dk_2\right)\int_{M_H^+} \delta(m)\chi_{z_0 z}(m)(1 - u(m, n_0))\, dm \\ &\quad + (1 + q^{-1})\left(\int_{K_H \times K_H} \phi_{z_0}^2(k_1, k_2)\, dk_1\, dk_2\right)\int_{M_H^+} \delta(m)\chi_{z_0^{-1} z}(m)(1 - u(m, n_0))\, dm\end{aligned}$$

is also well defined and has a meromorphic continuation at $z = 1$. Moreover this meromorphic continuation is holomorphic on $\mathcal{V} - \{1\}$ with at most a simple pole at $z_0 = 1$. As $u(\,\cdot\,, n_0)$ is compactly supported, we deduce that the integral

$$\begin{aligned}\int_H & \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))\, dh \\ &= \int_H \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))u(h, n_0)\, dh + \int_H \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))(1 - u(h, n_0))\, dh\end{aligned}$$

has a meromorphic continuation at $z = 1$ which we denote by

$$\int_H^* \varphi_{z_0}(h)\, dh.$$

The above discussion implies that $\int_H^* \varphi_{z_0}(h)\, dh$ is holomorphic on $\mathcal{V} - \{1\}$ with at most a simple pole at $z_0 = 1$.

The next result is established in [Feigon 2012, Proposition 4.6], but we think that the proof is not complete. We thank E. Lapid who suggested the proof below.

Proposition 5.1 (H-invariance). *For $x \in H$, we have*

$$\int_H^* \varphi_{z_0}(hx)\, dh = \int_H^* \varphi_{z_0}(h)\, dh.$$

Proof. We fix $x \in H$. For z, z' in $\mathbb{C}^*$, we set

$$F(\varphi_{z_0}, z, z')(h) := \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))\chi_{z'}(\mathcal{M}(hx^{-1})).$$

By (5-1), for $k_1, k_2 \in K_H$, and $m \in M_H^+$ with $h_{M_H}(m) > n_0$, we have

$$\delta_P(m)^{-1/2} F(\varphi_{z_0}, z, z')(k_1 m k_2) = \phi^1_{z_0}(k_1,k_2)\delta(m)(z_0 z)^{h_{M_H}(m)} z'^{h_{M_H}(\mathcal{M}(k_1 m k_2 x^{-1}))} + \phi^2_{z_0}(k_1,k_2)\delta(m)(z_0^{-1} z)^{h_{M_H}(m)} z'^{h_{M_H}(\mathcal{M}(k_1 m k_2 x^{-1}))}.$$

We can choose n_0 such that Lemma 2.1 is satisfied. Thus, for any $k_2 \in K_H$, there exists $X_{k_2 x^{-1}} \in \mathbb{R}$ such that, for any $m \in M_H^+$ satisfying $1 - u(m, n_0) \neq 0$, we have $h_{M_H}(\mathcal{M}(k_1 m k_2 x^{-1})) = h_{M_H}(m) + X_{k_2 x^{-1}}$. We deduce that

$$\begin{aligned}\delta_P(m)^{-1/2} & F(\varphi_{z_0}, z, z')(k_1 m k_2)(1 - u(m, n_0)) \\ &= \phi^1_{z_0}(k_1, k_2)\delta(m)(z_0 z z')^{h_{M_H}(m)} z'^{X_{k_2 x^{-1}}} + \phi^2_{z_0}(k_1, k_2)\delta(m)(z_0^{-1} z z')^{h_{M_H}(m)} z'^{X_{k_2 x^{-1}}}.\end{aligned}$$

Therefore, by Hartogs' theorem and the same argument as above, the function

$$(z_0, z, z') \mapsto \int_H \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))\chi_{z'}(\mathcal{M}(hx^{-1}))\, dh$$

is well defined for $|z_0 z z'| < 1$, and has a meromorphic continuation on $\mathcal{V} \times (\mathbb{C}^*)^2$. We denote by $I(\varphi_{z_0}, z, z')$ this meromorphic continuation. Moreover, for $z_0 \neq 1$, the function $(z, z') \mapsto I(\varphi_{z_0}, z, z')$ is holomorphic in a neighborhood of $(1, 1)$. For $|z_0 z| < 1$, we have $I(\varphi_{z_0}, z, 1) = \int_H \varphi_{z_0}(h)\chi_z(\mathcal{M}(h))\, dh$. Hence we deduce that

$$I(\varphi_{z_0}, 1, 1) = \int_H^* \varphi_{z_0}(h)\, dh.$$

On the other hand, we have $I(\varphi_{z_0}, 1, z') = \int_H \varphi_{z_0}(hx)\chi_{z'}(\mathcal{M}(h))\, dh$ for $|z_0 z'| < 1$, therefore, we obtain

$$I(\varphi_{z_0}, 1, 1) = \int_H^* \varphi_{z_0}(hx)\, dh.$$

This finishes the proof of the proposition. □

We will apply this to normalized Eisenstein integrals. Let $\delta \in \widehat{M}_2$ and $z \in \mathbb{C}^*$. Recall that we have defined the operator $C_{P,P}(1, \delta_z)$ by

$$C_{P,P}(1, \delta_z) := \mathrm{Id} \otimes A(\bar{P}, P, \check{\delta}_z) \in \mathrm{Hom}_G(i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}, i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}).$$

We set

$$\begin{aligned}C_{P,P}(w, \delta_z) &:= A(P, \bar{P}, w\delta_z)\lambda(w) \otimes \lambda(w) \\ &\in \mathrm{Hom}_G(i_P^G \mathbb{C}_{\delta_z} \otimes i_P^G \check{\mathbb{C}}_{\delta_z}, i_P^G \mathbb{C}_{w\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{w\delta_z}),\end{aligned}$$

where $\lambda(w)$ is the left translation by w which induces an isomorphism from $i_P^G \mathbb{C}_{\delta_z}$ to $i_{\bar{P}}^G \mathbb{C}_{w\delta_z}$. For $s \in W^G$, we define

$$\begin{aligned}C^0_{P,P}(s, \delta_z) &:= C_{P,P}(s, \delta_z) \circ C_{P,P}(1, \delta_z)^{-1} \\ &\in \mathrm{Hom}_G(i_P^G \mathbb{C}_{\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_z}, i_P^G \mathbb{C}_{s\delta_z} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{s\delta_z}).\end{aligned} \tag{5-2}$$

In particular, $C^0_{P,P}(1,\delta_z)$ is the identity map of $i^G_P\mathbb{C}_{\delta_z}\otimes i^G_{\overline{P}}\check{\mathbb{C}}_{\delta_z}$. By arguments analogous to those of [Waldspurger 2003, Lemme V.3.1], we obtain

(5-3) for $s\in W^G$ the rational operator $C^0_{P|P}(s,\delta_z)$ is regular on $\mathcal{O}$.

Let $S\in i^K_{P\cap K}\mathbb{C}\otimes i^K_{\overline{P}\cap K}\check{\mathbb{C}}$. By (3-7), the normalized Eisenstein integral $E^0(P,\delta_z,S_z)$ is holomorphic in a neighborhood $\mathcal{V}$ of $\mathcal{O}$. We may and will assume that $\mathcal{V}$ is invariant by the map $z\mapsto z^{-1}$. By [Heiermann 2001, Theorem 1.3.1] applied to $\lambda(k_1^{-1})\rho(k_2)E^0(P,\delta_z,S_z)$, $k_1,k_2\in K_H$, there exists a positive integer n_0 such that, for $k_1,k_2\in K_H$, and $m\in M_H^+$ satisfying $h_{M_H}(m)>n_0$, we have

$$\begin{aligned}&\delta_P(m)^{-1/2}E^0(P,\delta_z,S_z)(k_1mk_2)\\&=\delta(m)\big(\chi_z(m)\operatorname{tr}([C^0_{P,P}(1,\delta_z)S_z](k_1,k_2))+\chi_{z^{-1}}(m)\operatorname{tr}([C^0_{P,P}(w,\delta_z)S_z](k_1,k_2))\big).\end{aligned}$$

Therefore, the normalized Eisenstein integral satisfies (5-1). Hence, we can define the normalized regularized period by

$$P_{\delta_z}(S):=\int_H^* E^0(P,\delta_z,S_z)(h)\,dh,\quad S\in i^K_{P\cap K}\mathbb{C}\otimes i^K_{\overline{P}\cap K}\check{\mathbb{C}}. \tag{5-4}$$

The above discussion implies that $P_{\delta_z}(S)$ is a meromorphic function on the neighborhood $\mathcal{V}$ of $\mathcal{O}$ which is holomorphic on $\mathcal{V}-\{1\}$.

For $s\in W^G$ and $S\in i^K_{P\cap K}\mathbb{C}\otimes i^K_{\overline{P}\cap K}\check{\mathbb{C}}$, we set

$$C(s,\delta_z)(S):=(1+q^{-1})\int_{K_H\times K_H}\operatorname{tr}([C^0_{P,P}(s,\delta_z)S_z](k_1,k_2))\,dk_1\,dk_2. \tag{5-5}$$

By the same argument as in [Feigon 2012, Proposition 4.7], we have the following relations between the truncated period and the normalized regularized period.

(5-6) If $\delta_{|F^\times}\neq 1$ then, for n large enough, we have $P_{\delta_z}(S)=P^n_{\delta_z}(S)$.

(5-7) If $\delta_{|F^\times}=1$ then, for n large enough, we have

$$P_{\delta_z}(S)=P^n_{\delta_z}(S)+\frac{z^{n+1}}{1-z}C(1,\delta_z)(S)+\frac{z^{-(n+1)}}{1-z^{-1}}C(w,\delta_z)(S).$$

The following lemma is analogous to [Feigon 2012, Lemma 4.8].

Lemma 5.2. *Let $z\in\mathbb{C}^*$ and $S\in i^K_{P\cap K}\mathbb{C}\otimes i^K_{\overline{P}\cap K}\check{\mathbb{C}}$:*

(1) *If $\delta_{|F^\times}\neq 1$ and $\delta_{|E^1}\neq 1$ then, for n large enough, we have*

$$P_{\delta_z}(S)=P^n_{\delta_z}(S)=0.$$

(2) *If $\delta_{|F^\times}\neq 1$ and $\delta_{|E^1}=1$ then, for n large enough, we have*

$$P_{\delta_z}(S)=P^n_{\delta_z}(S).$$

(3) *If* $\delta_{|F^\times} = 1$ *and* $\delta_{|E^1} \neq 1$ *then* $P_{\delta_z}(S) = 0$ *whenever it is defined and*

$$C(1, \delta_1)(S) = C(w, \delta_1)(S).$$

(4) *If* $\delta_{|F^\times} = 1$ *and* $\delta_{|E^1} = 1$ *then* $\delta^2 = 1$. *We have* $C(1, \delta_1)(S) = -C(w, \delta_1)(S)$ *and the regularized normalized period* $P_{\delta_z}(S)$ *is meromorphic with a unique pole at* $z = 1$ *which is simple.*

Proof. Case (2) follows from (5-6). By [Jacquet et al. 1999, Proposition 22], if $\delta_{|E^1} \neq 1$ and $z \neq 1$ then the representation $i_P^G \delta_z$ admits no nontrivial H-invariant linear form. Thus in that case, Proposition 5.1 implies $P_{\delta_z}(S) = 0$ whenever it is defined. We deduce Case (1) from (5-6) and in Case (3), it follows from (5-7) that

$$P^n_{\delta_z}(S) = -\left(\frac{z^{n+1}}{1-z} C(1, \delta_z)(S) + \frac{z^{-(n+1)}}{1-z^{-1}} C(w, \delta_z)(S) \right).$$

Since $P^n_{\delta_z}(S)$ and $C(s, \delta_z)(S)$ for $s \in W^G$ are holomorphic functions at $z = 1$, and

$$\begin{aligned} &\operatorname{Res}\left(\frac{z^{n+1}}{1-z} C(1, \delta_z)(S), z = 1 \right) = -C(1, \delta_1)(S), \\ &\operatorname{Res}\left(\frac{z^{-(n+1)}}{1-z^{-1}} C(w, \delta_z)(S), z = 1 \right) = C(w, \delta_1)(S), \end{aligned} \tag{5-8}$$

we deduce the result in the Case (3).

In Case (4), we obtain easily $\delta^2 = 1$. By [Waldspurger 2003, Corollaire IV.1.2], the intertwining operator $A(\bar{P}, P, \delta_z)$ has a simple pole at $z = 1$. Thus the function $\mu(\delta_z)$ has a zero of order 2 at $z = 1$. In that case, by [Silberger 1979, proof of Theorem 5.4.2.1], the operators $C_{P|P}(s, \delta_z)$ for $s \in W^G$ have a simple pole at $z = 1$ and

$$\operatorname{Res}(C_{P|P}(1, \delta_z), z = 1) = -\operatorname{Res}(C_{P|P}(w, \delta_z), z = 1).$$

Therefore, if we set $T_z := (z-1)C_{P|P}(1, \delta_z)$ and $U_z := (z-1)C_{P|P}(w, \delta_z)$, then U_z and T_z^{-1} are holomorphic near $z = 1$ and $T_1 = -U_1$ as $\delta^2 = 1$. By definition (see (5-2)), we have $C^0_{P|P}(w, \delta_z) = U_z T_z^{-1}$. Hence, one deduces that $C^0_{P|P}(w, \delta_1) = -\operatorname{Id} = -C^0_{P|P}(1, \delta_1)$, where Id is the identity map of $i_P^G \mathbb{C}_{\delta_1} \otimes i_{\bar{P}}^G \check{\mathbb{C}}_{\delta_1}$. We deduce the first assertion in Case (4) from the definition of $C(s, \delta_z)(S)$ (see (5-5)).

Since $P^n_{\delta_z}(S)$ and $C(s, \delta_z)(S)$ for $s \in W^G$ are holomorphic functions at $z = 1$, the last assertion follows from (5-7), (5-8) and the above result. This finishes the proof of the lemma. □

6. A preliminary lemma

In this section, we prove a preliminary lemma which will allow us to get the asymptotic expansion of the truncated kernel in terms of regularized normalized periods.

Let $\mathcal{V}$ be a neighborhood of $\mathcal{O}$ in $\mathbb{C}^*$. We assume that $\mathcal{V}$ is invariant by the map $z \mapsto \bar{z}^{-1}$. Let f be a meromorphic function on $\mathcal{V}$. We assume that f has at most a pole at $z = 1$ in $\mathcal{V}$.

If $r \neq 1$ is such that f is defined on the set of complex numbers of modulus r, the integral $\int_{|z|=r} f(z)\,dz$ depends only on the position of r with respect to 1. We set

$$\int_{\mathcal{O}^-} f(z)\,dz := \int_{|z|=r} f(z)\,dz \quad \text{for } r < 1, \tag{6-1}$$

$$\int_{\mathcal{O}^+} f(z)\,dz := \int_{|z|=r} f(z)dz \quad \text{for } r > 1. \tag{6-2}$$

Notice that

$$\int_{\mathcal{O}^+} f(z)\,dz - \int_{\mathcal{O}^-} f(z)\,dz = 2i\pi \operatorname{Res}(f(z), z = 1). \tag{6-3}$$

It follows easily from the definitions that

$$\begin{aligned} &\lim_{n\to+\infty} \int_{\mathcal{O}^-} z^n f(z)\,dz = 0, \\ &\lim_{n\to+\infty} \int_{\mathcal{O}^+} z^{-n} f(z)\,dz = 0. \end{aligned} \tag{6-4}$$

We have assumed that $\mathcal{V}$ is invariant by the map $z \to \bar{z}^{-1}$. Then, the function $\tilde{f}(z) := \overline{f(\bar{z}^{-1})}$ is also a meromorphic function on $\mathcal{V}$ with at most a pole at $z = 1$ and it satisfies $\tilde{f}(z) = \overline{f(z)}$ for $z \in \mathcal{O}$.

Let $c(s, z)$ and $c'(s, z)$, for $s \in W^G$ be holomorphic functions on $\mathcal{V}$ such that $c(s, 1) \neq 0$ and $c'(s, 1) \neq 0$. Let p and p' be two meromorphic functions on $\mathcal{V}$ with at most a pole at $z = 1$. We set

$$\begin{aligned} p_n(z) &:= p(z) - \left[\frac{z^{n+1}}{1-z} c(1, z) + \frac{z^{-(n+1)}}{1 - z^{-1}} c(w, z)\right], \\ p'_n(z) &:= p'(z) - \left[\frac{z^{n+1}}{1-z} c'(1, z) + \frac{z^{-(n+1)}}{1-z^{-1}} c'(w, z)\right]. \end{aligned} \tag{6-5}$$

Lemma 6.1. *We assume that p_n and p'_n are holomorphic on $\mathcal{V}$ and that either p and p' are vanishing functions or $c(1, 1) = -c(w, 1)$ and $c'(1, 1) = -c'(w, 1)$. Then, the integral*

$$\int_{\mathcal{O}} p_n(z)\overline{p'_n(z)}\, \frac{dz}{z}$$

is asymptotic as n approaches $+\infty$ to the sum of

$$\int_{\mathcal{O}^-} \left(p(z)\tilde{p}'(z) + \frac{c(1, z)\tilde{c}'(1, z)}{(1 - z)(1 - z^{-1})} + \frac{c(w, z)\tilde{c}'(w, z)}{(1 - z)(1 - z^{-1})}\right) \frac{dz}{z}, \tag{6-6}$$

$$
\begin{aligned}
\text{(6-7)} \quad & -2i\pi\left[\frac{d}{dz}(c(w,z)\tilde{c}'(1,z))\right]_{z=1} \\
& \qquad + 2i\pi\left[\frac{d}{dz}\big(c(w,z)(z-1)\tilde{p}'(z)+\tilde{c}'(1,z)(z-1)p(z)\big)\right]_{z=1},
\end{aligned}
$$

and

$$
\begin{aligned}
\text{(6-8)} \quad & 2i\pi(2n+1)c(w,1)\tilde{c}'(1,1) \\
& \qquad - 2i\pi(n+1)\big(c(w,1)\operatorname{Res}(\tilde{p}',z=1)+\tilde{c}'(1,1)\operatorname{Res}(p,z=1)\big).
\end{aligned}
$$

Proof. Since p_n and $\tilde{p}_n'$ are holomorphic functions on $\mathcal{V}$, we have

$$
\begin{aligned}
\int_{\mathcal{O}} p_n(z)\overline{p_n'(z)}\,\frac{dz}{z} &= \int_{\mathcal{O}^-} p_n(z)\tilde{p}_n'(z)\,\frac{dz}{z} \\
&= \int_{\mathcal{O}^-} \Big(p(z) - \frac{z^{n+1}}{1-z}c(1,z) - \frac{z^{-(n+1)}}{1-z^{-1}}c(w,z)\Big) \\
&\qquad \times \Big(\tilde{p}'(z) - \frac{z^{-(n+1)}}{1-z^{-1}}\tilde{c}'(1,z) - \frac{z^{n+1}}{1-z}\tilde{c}'(w,z)\Big)\frac{dz}{z} \\
&= \int_{\mathcal{O}^-}\left(p(z)\tilde{p}'(z) + \frac{c(1,z)\tilde{c}'(1,z)}{(1-z)(1-z^{-1})} + \frac{c(w,z)\tilde{c}'(w,z)}{(1-z)(1-z^{-1})}\right)\frac{dz}{z} \\
&\qquad + \int_{\mathcal{O}^-} z^{2(n+1)}\frac{c(1,z)\tilde{c}'(w,z)}{(1-z)^2}\,\frac{dz}{z} \\
&\qquad - \int_{\mathcal{O}^-} z^{n+1}\left(\frac{c(1,z)\tilde{p}'(z)+p(z)\tilde{c}'(w,z)}{1-z}\right)\frac{dz}{z} \\
&\qquad + \int_{\mathcal{O}^-} z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{(1-z^{-1})^2}\,\frac{dz}{z} \\
&\qquad - \int_{\mathcal{O}^-} z^{-(n+1)}\left(\frac{c(w,z)\tilde{p}'(z)+p(z)\tilde{c}'(1,z)}{1-z^{-1}}\right)\frac{dz}{z}.
\end{aligned}
$$

By (6-4), the second and third terms of the right hand side converge to 0 as n approaches $+\infty$.

By (6-3), one has

$$
\begin{aligned}
&\int_{\mathcal{O}^-} z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{(1-z^{-1})^2}\,\frac{dz}{z} \\
&\quad = \int_{\mathcal{O}^+} z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{(1-z^{-1})^2}\,\frac{dz}{z} - 2i\pi\operatorname{Res}\left(z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{z(1-z^{-1})^2}, z=1\right).
\end{aligned}
$$

Let

$$
\phi(z) := z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{z(1-z^{-1})^2} = z^{-(2n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{(z-1)^2}.
$$

Since $c(w,z)$ and $\tilde{c}'(1,z)$ are holomorphic functions on $\mathcal{V}$, the function ϕ has a unique pole of order 2 at $z=1$. Thus, we obtain

$$\begin{aligned}\operatorname{Res}(\phi, z=1) &= \left[\frac{d}{dz}((z-1)^2\phi(z))\right]_{z=1}\\ &= -(2n+1)c(w,1)\tilde{c}'(1,1) + \left[\frac{d}{dz}(c(w,z)\tilde{c}'(1,z))\right]_{z=1}.\end{aligned}$$

We deduce from (6-4) that

$$\tag{6-9}\begin{aligned}&\int_{\mathcal{O}^-} z^{-2(n+1)}\frac{c(w,z)\tilde{c}'(1,z)}{(1-z^{-1})^2}\,\frac{dz}{z}\\ &\quad = 2i\pi(2n+1)c(w,1)\tilde{c}'(1,1) - 2i\pi\left[\frac{d}{dz}(c(w,z)\tilde{c}'(1,z))\right]_{z=1} + \epsilon_1(n),\end{aligned}$$

where $\lim_{n\to+\infty}\epsilon_1(n)=0$.

When p and p' are vanishing functions, we obtain the result of the lemma. Otherwise, by (6-5) and our assumptions, $(c(w,z)\tilde{p}'(z)+p(z)\tilde{c}'(1,z))/(1-z^{-1})$ is a meromorphic function with a unique pole of order 2 at $z=1$. Applying the same argument as above, we obtain

$$\begin{aligned}&\int_{\mathcal{O}^-} z^{-(n+1)}\left(\frac{c(w,z)\tilde{p}'(z)+p(z)\tilde{c}'(1,z)}{1-z^{-1}}\right)\frac{dz}{z}\\ &\quad= \int_{\mathcal{O}^+} z^{-(n+1)}\left(\frac{c(w,z)\tilde{p}'(z)+p(z)\tilde{c}'(1,z)}{1-z^{-1}}\right)\frac{dz}{z}\\ &\qquad - 2i\pi\left[\frac{d}{dz}\big(z^{-(n+1)}(z-1)(c(w,z)\tilde{p}'(z)+p(z)\tilde{c}'(1,z))\big)\right]_{z=1}\\ &\quad= 2i\pi(n+1)\big(c(w,1)\operatorname{Res}(\tilde{p}', z=1)+\operatorname{Res}(p,z=1)\tilde{c}'(1,1)\big)\\ &\qquad - 2i\pi\left[\frac{d}{dz}\big(c(w,z)(z-1)\tilde{p}'(z)+(z-1)p(z)\tilde{c}'(1,z)\big)\right]_{z=1} + \epsilon_2(n),\end{aligned}$$

where $\lim_{n\to+\infty}\epsilon_2(n)=0$.

Therefore, we obtain the lemma by (6-9) and the above result. □

7. Spectral side of a local relative trace formula

We recall the spectral expression of the truncated kernel obtained in Corollary 4.2:

$$\begin{aligned}K^n(f) &= \sum_{\tau\in\mathcal{E}_2(G)}\sum_{S\in\mathcal{B}_\tau} d(\tau)P^n_\tau(\tau\otimes\check{\tau}(f)S)\overline{P^n_\tau(S)}\\ &\qquad + \frac{1}{4i\pi}\sum_{\delta\in\widehat{M}_2}\sum_{S\in\mathcal{B}_{\bar{P},P}(E)}\int_{\mathcal{O}} P^n_{\delta_z}(\overline{\Pi}_{\delta_z}(f)S)\overline{P^n_{\delta_z}(S)}\,\frac{dz}{z},\end{aligned}$$

where the sums over S are all finite and $\overline{\Pi}_{\delta_z} := \bar{i}_P^G \delta_z \otimes \bar{i}_{\bar{P}}^G \check{\delta}_z$.

By [Feigon 2012, Lemma 4.10], if $(\tau, V_\tau) \in \mathcal{E}_2(G)$ and $S \in \mathrm{End}_{fin.rk}(V_\tau)$, then

$$\lim_{n\to+\infty} P_\tau^n(S) = \int_H \operatorname{tr}(\tau(h)S)\, dh. \tag{7-1}$$

We consider now the second term of the above expression of $K^n(f)$. Let $\delta \in \widehat{M}_2$ and $S \in i_{P\cap K}^K \mathbb{C} \otimes i_{\bar{P}\cap K}^K \check{\mathbb{C}}$. We keep the notation of the previous section. In particular, for $z \in \mathbb{C}^*$, we have $\widetilde{C}(s, \delta_z)(S) = \overline{C(s, \delta_{\bar{z}^{-1}})(S)}$ and $\widetilde{P}_{\delta_z}(S) = \overline{P_{\delta_{\bar{z}^{-1}}}(S)}$. By the definition of δ_z, we have $\delta_1 = \delta$.

Proposition 7.1. *Let $S \in i_{P\cap K}^K \mathbb{C} \otimes i_{\bar{P}\cap K}^K \check{\mathbb{C}}$. We set $S'_z := \overline{\Pi}_{\delta_z}(f)S$:*

(1) *If $\delta_{|F^\times} \neq 1$ and $\delta_{|E^1} \neq 1$ then, for $n \in \mathbb{N}$ large enough, one has*

$$\int_{\mathcal{O}} P_{\delta_z}^n(S'_z)\,\overline{P_{\delta_z}^n(S)}\,\frac{dz}{z} = 0.$$

(2) *If $\delta_{|F^\times} \neq 1$ and $\delta_{|E^1} = 1$ then*

$$\lim_{n\to+\infty} \int_{\mathcal{O}} P_{\delta_z}^n(S'_z)\,\overline{P_{\delta_z}^n(S)}\,\frac{dz}{z} = \int_{\mathcal{O}} P_{\delta_z}(S'_z)\,\overline{P_{\delta_z}(S)}\,\frac{dz}{z}.$$

(3) *Assume that $\delta_{|F^\times} = 1$ and $\delta_{|E^1} \neq 1$. Then*

$$\int_{\mathcal{O}} P_{\delta_z}^n(S'_z)\,\overline{P_{\delta_z}^n(S)}\,\frac{dz}{z}$$

is asymptotic when n approaches $+\infty$ to

$$\begin{aligned}
&2i\pi(2n+1)C(1,\delta)(S'_1)\overline{C(1,\delta)(S)} \\
&\quad + \int_{\mathcal{O}^-} \left(\frac{C(1,\delta_z)(S'_z)\widetilde{C}(1,\delta_z)(S)}{(1-z)(1-z^{-1})} + \frac{C(w,\delta_z)(S'_z)\widetilde{C}(w,\delta_z)(S)}{(1-z)(1-z^{-1})} \right) \frac{dz}{z} \\
&\qquad\qquad - 2i\pi \frac{d}{dz}[C(w,\delta_z)(S'_z)\widetilde{C}(1,\delta_z)(S)]_{z=1}.
\end{aligned}$$

(4) *Assume that $\delta_{|F^\times} = 1$ and $\delta_{|E^1} = 1$. Then*

$$\int_{\mathcal{O}} P_{\delta_z}^n(S'_z)\,\overline{P_{\delta_z}^n(S)}\,\frac{dz}{z}$$

is asymptotic when n approaches $+\infty$ to

$$\begin{aligned}
&2i\pi(2n+3)C(1,\delta)(S'_1)\overline{C(1,\delta)(S)} \\
&\quad + \int_{\mathcal{O}^-} \left(P_{\delta_z}(S'_z)\overline{P_{\delta_z}(S)} + \frac{C(1,\delta_z)(S'_z)\widetilde{C}(1,\delta_z)(S)}{(1-z)(1-z^{-1})} + \frac{C(w,\delta_z)(S'_z)\widetilde{C}(w,\delta_z)(S)}{(1-z)(1-z^{-1})} \right) \frac{dz}{z} \\
&\quad - 2i\pi \frac{d}{dz}[C(w,\delta_z)(S'_z)\widetilde{C}(1,\delta_z)(S)]_{z=1} \\
&\qquad + 2i\pi \left[\frac{d}{dz}\big((z-1)P_{\delta_z}(S'_z)\widetilde{C}(1,\delta_z)(S) + C(w,\delta_z)(S'_z)(z-1)\widetilde{P}_{\delta_z}(S)\big) \right]_{z=1}.
\end{aligned}$$

Proof. The two first assertions are immediate consequences of Lemma 5.2. To prove (3) and (4), we set

$$p_n(z) := P^n_{\delta_z}(S'_z(f)), \quad p'_n(z) := P^n_{\delta_z}(S), \quad p(z) := P_{\delta_z}(S'_z(f)), \quad p'(z) := P_{\delta_z}(S)$$
$$\text{and} \quad c(s,z) := C(s,\delta_z)(S'_z(f)), \quad c'(s,z) := C(s,\delta_z)(S), \quad \text{for } s \in W^G.$$

By (5-7) and Lemma 5.2 these functions satisfy (6-5) and we can apply Lemma 6.1. The result in case (3) follows immediately since $p(z) = p'(z) = 0$ by Lemma 5.2.

In case (4), we have $c(1,1) = -c(w,1)$ and $c'(1,1) = -c'(w,1)$ by Lemma 5.2. Moreover, the relations (6-5) give

$$\mathrm{Res}(p, z=1) = -c(1,1) + c(w,1) \quad \text{and} \quad \mathrm{Res}(\tilde{p}', z=1) = c'(1,1) - c'(w,1).$$

Hence, we obtain

$$2i\pi(2n+1)c(w,1)\tilde{c}'(1,1) - 2i\pi(n+1)\big(c(w,1)\mathrm{Res}(\tilde{p}', z=1) + \tilde{c}'(1,1)\mathrm{Res}(p, z=1)\big).$$
$$= 2i\pi(2n+3)c(1,1)\tilde{c}'(1,1),$$

and the result in that case follows from Lemma 6.1. □

To describe the spectral side of our local relative trace formula, we introduce generalized matrix coefficients.

Let (π, V) be a smooth unitary representation of G. We denote by (π', V') its dual representation. Let ξ and ξ' be two linear forms on V. For $f \in C_c^\infty(G)$, the linear form $\pi'(\check{f})\xi$ belongs to the smooth dual $\check{V}$ of V [Renard 2010, Théorème III.3.4 and I.1.2]. The scalar product on V induces an isomorphism $j : v \mapsto (\cdot, v)$ from the conjugate complex vector space $\overline{V}$ of V and $\check{V}$, which intertwines the complex conjugate of π and $\check{\pi}$ as π is unitary. One has

$$\check{v}(v) = (v, j^{-1}(\check{v})), \quad v \in V, \check{v} \in \check{V}.$$

Therefore, for $v \in V$, we have

$$(\pi'(\check{f})\xi)(v) = \xi(\pi(f)v) = (v, j^{-1}(\pi'(\check{f})\xi)).$$

As $\pi(f)$ is an operator of finite rank, we have for any orthonormal basis $\mathcal{B}$ of V

$$\text{(7-2)} \qquad j^{-1}(\pi'(\check{f})\xi) = \sum_{v \in \mathcal{B}} (\pi'(\check{f})\xi)(v) \cdot v,$$

where the sum over v is finite, and $(\lambda, v) \mapsto \lambda \cdot v$ is the action of $\mathbb{C}$ on $\overline{V}$.

Let $\overline{\xi'}$ be the linear form on $\overline{V}$ defined by $\overline{\xi'}(u) = \overline{\xi'(u)}$. We define the generalized matrix coefficient $m_{\xi,\xi'}$ by

$$m_{\xi,\xi'}(f) = \overline{\xi'}(j^{-1}(\pi'(\check{f})\xi)).$$

Then, by (7-2), we obtain

$$m_{\xi,\xi'}(f) = \sum_{v\in\mathcal{B}} \xi(\pi(f)v)\overline{\xi'}(v). \tag{7-3}$$

Hence, this sum is independent of the choice of the basis $\mathcal{B}$.

Let $z \in \mathbb{C}^*$. We set $(\Pi_z, V_z) := (i_P^G \delta_z \otimes i_{\overline{P}}^G \check{\delta_{\delta_z}}, i_P^G \mathbb{C}_{\delta_z} \otimes i_{\overline{P}}^G \check{\mathbb{C}_{\delta_z}})$. We denote by $(\overline{\Pi}_z, V)$ its compact realization. We define meromorphic linear forms on V_z using the isomorphism $V_z \simeq V$.

Lemma 7.2. *Let ξ_z and ξ'_z be two linear forms on V which are meromorphic in z on a neighborhood $\mathcal{V}$ of $\mathcal{O}$. Let $\mathcal{B}$ be an orthonormal basis of V. Then, for $f \in C_c^\infty(G\times G)$, the sum*

$$\sum_{S\in\mathcal{B}} \xi_z(\overline{\Pi}_z(f)S)\overline{\xi_{\bar{z}^{-1}}(S)}$$

is a finite sum over S which is independent of the choice of the basis $\mathcal{B}$.

Proof. For $z \in \mathcal{O}$, the representation Π_z is unitary. Hence (7-3) gives the lemma in that case. Since the linear forms ξ_z and ξ'_z are meromorphic on $\mathcal{V}$, we deduce the result of the lemma for any z in $\mathcal{V}$ by meromorphic continuation. □

With notation of the lemma, we define, for $z \in \mathcal{V}$, the generalized matrix coefficient $m_{\xi_z,\xi'_{\bar{z}^{-1}}}$ associated to (ξ_z, ξ'_z) by

$$m_{\xi_z,\xi'_{\bar{z}^{-1}}}(f) := \sum_{S\in\mathcal{B}} \xi_z(\overline{\Pi}_z(f)S)\overline{\xi_{\bar{z}^{-1}}(S)}.$$

Therefore, using Proposition 7.1, we can deduce the asymptotic behavior of the truncated kernel in terms of generalized matrix coefficients.

Theorem 7.3. *As n approaches $+\infty$, the truncated kernel $K^n(f)$ is asymptotic to*

$$n \sum_{\substack{\delta\in\widehat{M}_2\\ \delta_{|F^\times}=1}} m_{C(1,\delta),C(1,\delta)}(f) + \sum_{\tau\in\mathcal{E}_2(G)} d(\tau) m_{P_\tau,P_\tau}(f) + \frac{1}{4i\pi} \sum_{\substack{\delta\in\widehat{M}_2\\ \delta_{|F^\times}\neq 1\\ \delta_{|E^1}=1}} \int_{\mathcal{O}} m_{P_{\delta_z},P_{\delta_z}}(f)\,\frac{dz}{z}$$

$$+ \frac{1}{4i\pi} \sum_{\substack{\delta\in\widehat{M}_2\\ \delta_{|F^\times}=1\\ \delta_{|E^1}\neq 1}} \left(R_\delta(f) + \int_{\mathcal{O}^-} \frac{m_{C(1,\delta_z),C(1,\delta_{\bar{z}^{-1}})}(f) + m_{C(w,\delta_z),C(w,\delta_{\bar{z}^{-1}})}(f)}{(1-z)(1-z^{-1})}\,\frac{dz}{z}\right)$$

$$+ \frac{1}{4i\pi} \sum_{\substack{\delta\in\widehat{M}_2\\ \delta_{|F^\times}=\delta_{|E^1}=1}} \left(\widetilde{R}_\delta(f) + \int_{\mathcal{O}^-} \frac{m_{C(1,\delta_z),C(1,\delta_{\bar{z}^{-1}})}(f) + m_{C(w,\delta_z),C(w,\delta_{\bar{z}^{-1}})}(f)}{(1-z)(1-z^{-1})}\,\frac{dz}{z}\right.$$

$$\left. + \int_{\mathcal{O}^-} m_{P_{\delta_z},P_{\delta_{\bar{z}^{-1}}}}(f)\,\frac{dz}{z}\right).$$

where

$$R_\delta(f) := 2i\pi\left(m_{C(1,\delta),C(1,\delta)}(f) - \left[\frac{d}{dz} m_{C(w,\delta_z),C(1,\delta_{\bar{z}^{-1}})}(f)\right]_{z=1}\right),$$

$$\widetilde{R}_\delta(f) = 2i\pi\Bigg(3m_{C(1,\delta),C(1,\delta)}(f) - \left[\frac{d}{dz} m_{C(w,\delta_z),C(1,\delta_{\bar{z}^{-1}})}(f)\right]_{z=1} + \left[\frac{d}{dz}(z-1)\big(m_{P_{\delta_z},C(1,\delta_{\bar{z}^{-1}})}(f) + m_{C(w,\delta_z),P_{\delta_{\bar{z}^{-1}}}}(f)\big)\right]_{z=1}\Bigg),$$

$$P_\tau(S) = \int_H \operatorname{tr}(\tau(h)S)\,dh, \quad S \in \operatorname{End}_{fin.rk}(V_\tau),$$

$$P_{\delta_z}(S) = \int_H^* E^0(P,\delta_z,S_z)(h)\,dh, \quad S \in i_{P\cap K}^K \mathbb{C} \otimes i_{\bar{P}\cap K}^K \check{\mathbb{C}}$$

$$C(s,\delta_z)(S) := (1+q^{-1}) \int_{K_H\times K_H} \operatorname{tr}([C^0_{P,P}(s,\delta_z)S_z](k_1,k_2))\,dk_1\,dk_2, \quad s \in W^G.$$

8. A local relative trace formula for PGL(2)

We make precise the geometric expansion of the truncated kernel obtained in [Delorme et al. 2015, Theorem 2.3] for $\underline{H} := \mathrm{PGL}(2)$. This geometric expansion depends on orbital integrals of f_1 and f_2, and on a weight function v_L where $L = H$ or M. To recall the definition of these objects, we need to introduce some notation.

If $\underline{J}$ is an algebraic group defined over F, we denote by J its group of points over F and we identify $\underline{J}$ with the group of points of $\underline{J}$ over an algebraic closure of F. Let $\underline{J}_H$ be an algebraic subgroup of $\underline{H}$ defined over F. We denote by $\underline{J} := \mathrm{Res}_{E/F}(\underline{J}_H \times_F E)$ the restriction of scalars of $\underline{J}_H$ from E to F. Then, the group $J := \underline{J}(F)$ is isomorphic to $\underline{J}_H(E)$.

The nontrivial element of the Galois group of E/F induces an involution σ of $\underline{G}$ defined over F.

We denote by $\underline{\mathcal{P}}$ the connected component of 1 in the set of x in $\underline{G}$ such that $\sigma(x) = x^{-1}$. A torus $\underline{A}$ of $\underline{G}$ is called a σ-torus if $\underline{A}$ is a torus defined over F contained in $\underline{\mathcal{P}}$. Let $\underline{S}_H$ be a maximal torus of $\underline{H}$. We denote by $\underline{S}_\sigma$ the connected component of $\underline{S} \cap \underline{\mathcal{P}}$. Then $\underline{S}_\sigma$ is a maximal σ-torus defined over F and the map $S_H \mapsto S_\sigma$ is a bijective correspondence between H-conjugacy classes of maximal tori of H and H-conjugacy classes of maximal σ-tori of G (see [Delorme et al. 2015, 1.2]).

Each maximal torus of H is either anisotropic or H-conjugate to M. We fix $\mathcal{T}_H$ a set of representatives for the H-conjugacy classes of maximal anisotropic torus in H.

By [ibid., 1.28], for each maximal torus S_H of H, we can fix a finite set of representatives $\kappa_S = \{x_m\}$ of the (H, S_σ)-double cosets in $\underline{H}\,\underline{S}_\sigma \cap G$ such that each element x_m may be written $x_m = h_m a_m^{-1}$ where $h_m \in \underline{H}$ centralizes the split component A_S of S_H and $a_m \in \underline{S}_\sigma$.

The orbital integral of a compactly supported smooth function is defined on the set $G^{\sigma\text{-reg}}$ of σ-regular points of G, that is the set of point x in G such that $\underline{H}x\underline{H}$ is Zariski closed and of maximal dimension. The set $G^{\sigma\text{-reg}}$ can be described in terms of maximal σ-tori as follows. If $\underline{S}_H$ is a maximal torus of $\underline{H}$, we denote by $\underline{\mathfrak{s}}$ the Lie algebra of $\underline{S}$ and we set $\mathfrak{s} := \underline{\mathfrak{s}}(F)$. We set

$$\Delta_\sigma(g) = \det(1 - \operatorname{Ad}(g^{-1}\sigma(g))_{\mathfrak{g}/\mathfrak{s}}), \quad g \in G.$$

By [Delorme et al. 2015, 1.30], if $x \in G^{\sigma\text{-reg}}$ then there exists a maximal torus S_H of H such that $\Delta_\sigma(x) \neq 0$. Moreover, there are two elements $x_m \in \kappa_S$ and $\gamma \in S_\sigma$ such that $x = x_m\gamma$.

We define the orbital integral $\mathcal{M}(f)$ of a function $f \in C_c^\infty(G)$ on $G^{\sigma\text{-reg}}$ as follows. Let S_H be a maximal torus of H. For $x_m \in \kappa_S$ and $\gamma \in S_\sigma$ with $\Delta_\sigma(x_m\gamma) \neq 0$, we set

$$\text{(8-1)} \qquad \mathcal{M}(f)(x_m\gamma) := |\Delta_\sigma(x_m\gamma)|_F^{1/4} \int_{\operatorname{diag}(A_S)\backslash(H\times H)} f(h^{-1}x_m\gamma l)\, d\overline{(h,l)},$$

where $\operatorname{diag}(A_S)$ is the diagonal of $A_S \times A_S$.

We now give an explicit expression of the truncated function $v_L(\,\cdot\,, n)$ defined in [ibid., 2.12], where n is a positive integer and L is equal to H or M. Let n be a positive integer. It follows immediately from the definition [ibid., 2.12] that we have

$$\text{(8-2)} \qquad v_H(x_1, y_1, x_2, y_2, n) = 1, \quad x_1, y_1, x_2, y_2 \in H.$$

We will describe v_M using [ibid., 2.6]. Since $H = P_H K_H$, each $x \in H$ can be written $x = m_{P_H}(x)n_{P_H}(x)k_{P_H}(x)$ with $m_{P_H}(x) \in M_H$, $n_{P_H}(x) \in N_H$ and $k_{P_H}(x) \in K_H$. We take similar notation if we consider $\bar{P}$ instead of P. For $Q = P$ or $\bar{P}$, we set

$$h_{Q_H}(x) := h_{M_H}(m_{Q_H}(x)).$$

With our definition of h_{M_H} (2-2), the map $M_H \to \mathbb{R}$ given in [ibid., 1.2] coincides with $-(\log q)h_{M_H}$.

For x_1, y_1, x_2 and y_2 in H, we set

$$\begin{aligned} z_P(x_1, y_1, x_2, y_2) &:= \inf(h_{\bar{P}_H}(x_1) - h_{P_H}(y_1), h_{\bar{P}_H}(x_2) - h_{P_H}(y_2)), \quad \text{and} \\ z_{\bar{P}}(x_1, y_1, x_2, y_2) &:= -\inf(h_{\bar{P}_H}(y_1) - h_{P_H}(x_1), h_{\bar{P}_H}(y_2) - h_{P_H}(x_2)). \end{aligned}$$

We omit x_1, y_1, x_2 and y_2 in this notation if there is no confusion. Hence the elements Z_P^0 and $Z_{\bar{P}}^0$ of [ibid., 2.55] coincide with $(\log q)z_P$ and $(\log q)z_{\bar{P}}$ respectively. Therefore, the relation [ibid., 2.63] gives

$$\begin{aligned} v_M(x_1, y_1, x_2, y_2, n) &= \lim_{\lambda\to 0}\left(\frac{q^{\lambda(n+z_P)}}{1-q^{-2\lambda}}(1+q^{-\lambda}) + \frac{q^{\lambda(-n+z_{\bar{P}})}}{1-q^{2\lambda}}(1+q^{\lambda})\right) \\ &= \lim_{\lambda\to 0}\left(\frac{q^{\lambda(n+z_P)}}{1-q^{-\lambda}} + \frac{q^{-\lambda(n-z_{\bar{P}})}}{1-q^{\lambda}}\right) \\ &= \lim_{\lambda\to 0}\frac{q^{\lambda(n+z_P)} - q^{-\lambda(n-z_{\bar{P}}+1)}}{1-q^{-\lambda}} \\ &= 2n+1+z_P - z_{\bar{P}}. \end{aligned}$$

We set

$$\begin{aligned} v_M^0(x_1, y_1, x_2, y_2) &:= z_P - z_{\bar{P}} \\ &= \inf(h_{\bar{P}_H}(x_1) - h_{P_H}(y_1), h_{\bar{P}_H}(x_2) - h_{P_H}(y_2)) \\ &\quad + \inf(h_{\bar{P}_H}(y_1) - h_{P_H}(x_1), h_{\bar{P}_H}(y_2) - h_{P_H}(x_2)). \end{aligned}$$

Therefore, [ibid., Theorem 2.3] gives that as n approaches $+\infty$, the truncated kernel $K^n(f)$ is asymptotic to

$$\begin{aligned} \text{(8-3)}\quad 2n \sum_{x_m\in\kappa_M} c^0_{M,x_m} \int_{M_\sigma} \mathcal{M}(f_1)(x_m\gamma)\mathcal{M}(f_2)(x_m\gamma)\,d\gamma \\ + \sum_{S_H\in\mathcal{T}_H\cup\{M_H\}} \sum_{x_m\in\kappa_S} c^0_{S,x_m} \int_{S_\sigma} \mathcal{M}(f_1)(x_m\gamma)\mathcal{M}(f_2)(x_m\gamma)\,d\gamma \\ + \sum_{x_m\in\kappa_M} c^0_{M,x_m} \int_{M_\sigma} \mathcal{W}\mathcal{M}(f)(x_m\gamma)\,d\gamma, \end{aligned}$$

where the constants c^0_{M,x_m} are defined in [Rader and Rallis 1996, Theorem 3.4] and $\mathcal{W}\mathcal{M}(f)$ is the weighted integral orbital given by

$$\begin{aligned} &\Delta_\sigma(x_m\gamma)^{-1/2}\mathcal{W}\mathcal{M}(f)(x_m\gamma) \\ &= \iint_{(\mathrm{diag}(M_H)\backslash H\times H)^2} f_1(x_1^{-1}x_m\gamma x_2) f_2(y_1^{-1}x_m\gamma y_2) v_M^0(x_1, y_1, x_2, y_2)\, d\overline{(x_1, x_2)}\, d\overline{(y_1, y_2)}. \end{aligned}$$

Therefore, comparing asymptotic expansions of $K^n(f)$ in Theorem 7.3 and (8-3), we obtain:

Theorem 8.1. *For f_1 and f_2 in $C_c^\infty(G)$ we have*:

$$\text{(1)}\quad 2\sum_{x_m\in\kappa_M} c^0_{M,x_m} \int_{M_\sigma} \mathcal{M}(f_1)(x_m\gamma)\mathcal{M}(f_2)(x_m\gamma)\,d\gamma = \sum_{\substack{\delta\in\widehat{M}_2 \\ \delta_{|F^\times}=1}} m_{C(1,\delta),C(1,\delta)}(f).$$

(2) (Local relative trace formula) *The expression*

$$\sum_{S_H \in \mathcal{T}_H \cup \{M_H\}} \sum_{x_m \in \kappa_S} c^0_{S,x_m} \int_{S_\sigma} \mathcal{M}(f_1)(x_m\gamma)\mathcal{M}(f_2)(x_m\gamma)\, d\gamma + \sum_{x_m \in \kappa_M} c^0_{M,x_m} \int_{M_\sigma} \mathcal{W}\mathcal{M}(f)(x_m\gamma)\, d\gamma$$

equals

$$\sum_{\tau \in \mathcal{E}_2(G)} d(\tau) m_{P_\tau,P_\tau}(f) + \frac{1}{4i\pi} \sum_{\substack{\delta \in \widehat{M}_2 \\ \delta_{|F^\times} \neq 1 \\ \delta_{|E^1}=1}} \int_{\mathcal{O}} m_{P_{\delta_z},P_{\delta_z}}(f)\, \frac{dz}{z}$$
$$+ \frac{1}{4i\pi} \sum_{\substack{\delta \in \widehat{M}_2 \\ \delta_{|F^\times}=1 \\ \delta_{|E^1}\neq 1}} \left(R_\delta(f) + \int_{\mathcal{O}^-} \frac{m_{C(1,\delta_z),C(1,\delta_{\bar z^{-1}})}(f) + m_{C(w,\delta_z),C(w,\delta_{\bar z^{-1}})}(f)}{(1-z)(1-z^{-1})}\, \frac{dz}{z} \right)$$
$$+ \frac{1}{4i\pi} \sum_{\substack{\delta \in \widehat{M}_2 \\ \delta_{|F^\times}=\delta_{|E^1}=1}} \left(\widetilde{R}_\delta(f) + \int_{\mathcal{O}^-} \frac{m_{C(1,\delta_z),C(1,\delta_{\bar z^{-1}})}(f) + m_{C(w,\delta_z),C(w,\delta_{\bar z^{-1}})}(f)}{(1-z)(1-z^{-1})}\, \frac{dz}{z} + \int_{\mathcal{O}^-} m_{P_{\delta_z},P_{\delta_{\bar z^{-1}}}}(f)\, \frac{dz}{z} \right),$$

where

$$\begin{aligned}
R_\delta(f) &:= 2i\pi \left(m_{C(1,\delta),C(1,\delta)}(f) - \left[\frac{d}{dz} m_{C(w,\delta_z),C(1,\delta_{\bar z^{-1}})}(f) \right]_{z=1} \right), \\
\widetilde{R}_\delta(f) &= 2i\pi \Bigg(3m_{C(1,\delta),C(1,\delta)}(f) - \left[\frac{d}{dz} m_{C(w,\delta_z),C(1,\delta_{\bar z^{-1}})}(f) \right]_{z=1} \\
&\qquad + \left[\frac{d}{dz}(z-1)\big(m_{P_{\delta_z},C(1,\delta_{\bar z^{-1}})}(f) + m_{C(w,\delta_z),P_{\delta_{\bar z^{-1}}}}(f)\big) \right]_{z=1} \Bigg), \\
P_\tau(S) &= \int_{H_*} \operatorname{tr}(\tau(h)S)\, dh, \quad \text{for } S \in \operatorname{End}(V_\tau), \\
P_{\delta_z}(S) &= \int_H E^0(P, \delta_z, S_z)(h)\, dh, \quad \text{for } S \in i^K_{P\cap K}\mathbb{C} \otimes i^K_{\bar P\cap K}\check{\mathbb{C}}, \\
C(s,\delta_z)(S) &:= (1+q^{-1}) \int_{K_H \times K_H} \operatorname{tr}([C^0_{\bar P,P}(s,\delta_z)S_z](k_1,k_2))\, dk_1\, dk_2, \quad \text{for } s \in W^G.
\end{aligned}$$

As an application of this theorem, we will invert orbital integrals on the anisotropic σ-torus M_σ of G.

Let $\delta \in \widehat{M}_2$. As the operator $C^0_{P,P}(1,\delta)$ is the identity operator of $i^G_P \mathbb{C}_{\delta_z} \otimes i^G_{\bar P} \check{\mathbb{C}}_{\delta_z}$, one has

$$C(1,\delta)(v\otimes\check{w})=(1+q^{-1})\int_{K_H\times K_H} v(k_1)\check{w}(k_2)\,dk_1\,dk_2, \quad v\otimes\check{w}\in i_{P\cap K}^K\mathbb{C}\otimes i_{\bar{P}\cap K}^K\check{\mathbb{C}}.$$

Hence, we have $C(1,\delta)=(1+q^{-1})\xi_\delta\otimes\xi_{\check{\delta}}$ where ξ_δ and $\xi_{\check{\delta}}$ are the H-invariant linear forms on $i_{P\cap K}^K\mathbb{C}$ and $i_{\bar{P}\cap K}^K\check{\mathbb{C}}$ respectively given by the integration over K_H. Therefore, one deduces that

$$m_{C(1,\delta),C(1,\delta)}(f_1\otimes f_2)=m_{\xi_\delta,\xi_\delta}(f_1)m_{\xi_{\check{\delta}},\xi_{\check{\delta}}}(f_2).$$

Moreover, by [Aizenbud et al. 2015, Corollary 5.6.3], the distribution $f\mapsto m_{\xi_{\check{\delta}},\xi_{\check{\delta}}}(f)$ is smooth in a neighborhood of any σ-regular point of G.

Corollary 8.2. *Let $f\in C_c^\infty(G)$. Let $x_m\in\kappa_M$ and $\gamma\in M_\sigma$ such that $x_m\gamma$ is σ-regular. Then we have*

$$c^0_{M,x_m}|\Delta_\sigma(x_m\gamma)|^{1/4}\mathcal{M}(f)(x_m\gamma)=\sum_{\delta\in\widehat{M}_2,\delta_{|F^\times}=1} m_{\xi_\delta,\xi_\delta}(f)m_{\xi_{\check{\delta}},\xi_{\check{\delta}}}(x_m\gamma).$$

Proof. Let $(J_n)_n$ be a sequence of compact open subgroups whose intersection is equal to the neutral element of G. Then the characteristic function g_n of $J_nx_m\gamma J_n$ approaches the Dirac measure at $x_m\gamma$. Therefore, taking $f_1:=f$ and $f_2:=g_n$ in Theorem 8.1(1), we obtain the result. □

Remark. Let (τ, V_τ) be a supercuspidal representation of G and f be a matrix coefficient of τ. Then we deduce from the corollary that the orbital integral of f on σ-regular points of M_σ is equal to 0.

We assume that (τ, V_τ) is H-distinguished. By [Flicker 1991, Proposition 11] we have $\dim V_\tau'^H=1$. Let ξ be a nonzero H-invariant linear form on V_τ. Let S_H be an anisotropic torus of H and $x_m\in\kappa_S$. Then, applying our local relative trace formula to $f_1:=f$ and f_2 approaching the Dirac measure at a σ-regular point $x_m\gamma$ with $\gamma\in S_\sigma$, we obtain

$$|\Delta_\sigma(x_m\gamma)|^{1/4}\mathcal{M}(f)(x_m\gamma)=cm_{\xi,\xi}(f)m_{\xi,\xi}(x_m\gamma),$$

where c is some nonzero constant.

J. Hakim [1991, Proposition 8.1 and Lemma 8.1] obtained these results by other methods.

Acknowledgments

We thank the referee for his useful comments.

References

[Aizenbud et al. 2015] A. Aizenbud, D. Gourevitch, and E. Sayag, "𝔷-finite distributions on p-adic groups", *Adv. Math.* **285** (2015), 1376–1414. MR Zbl

[Arthur 1991] J. Arthur, "A local trace formula", *Inst. Hautes Études Sci. Publ. Math.* 73 (1991), 5–96. MR Zbl

[Delorme et al. 2015] P. Delorme, P. Harinck, and S. Souaifi, "Geometric side of a local relative trace formula", preprint, 2015. To appear in *Trans. Amer. Math. Soc.* arXiv

[Feigon 2012] B. Feigon, "A relative trace formula for PGL(2) in the local setting", *Pacific J. Math.* **260**:2 (2012), 395–432. MR Zbl

[Flicker 1991] Y. Z. Flicker, "On distinguished representations", *J. Reine Angew. Math.* **418** (1991), 139–172. MR Zbl

[Hakim 1991] J. Hakim, "Distinguished p-adic representations", *Duke Math. J.* **62**:1 (1991), 1–22. MR Zbl

[Heiermann 2001] V. Heiermann, "Une formule de Plancherel pour l'algèbre de Hecke d'un groupe réductif p-adique", *Comment. Math. Helv.* **76**:3 (2001), 388–415. MR Zbl

[Jacquet et al. 1999] H. Jacquet, E. Lapid, and J. Rogawski, "Periods of automorphic forms", *J. Amer. Math. Soc.* **12**:1 (1999), 173–240. MR Zbl

[Rader and Rallis 1996] C. Rader and S. Rallis, "Spherical characters on p-adic symmetric spaces", *Amer. J. Math.* **118**:1 (1996), 91–178. MR Zbl

[Renard 2010] D. Renard, *Représentations des groupes réductifs p-adiques*, Cours Spécialisés **17**, Société Mathématique de France, Paris, 2010. MR Zbl

[Silberger 1979] A. J. Silberger, *Introduction to harmonic analysis on reductive p-adic groups*, Mathematical Notes **23**, Princeton Univ. Press, 1979. MR Zbl

[Waldspurger 2003] J.-L. Waldspurger, "La formule de Plancherel pour les groupes p-adiques (d'après Harish-Chandra)", *J. Inst. Math. Jussieu* **2**:2 (2003), 235–333. MR Zbl

Received September 19, 2016. Revised March 24, 2017.

PATRICK DELORME
CNRS, CENTRALE MARSEILLE, I2M, UMR 7373
AIX-MARSEILLE UNIVERSITÉ
13453 MARSEILLE
FRANCE
patrick.delorme@univ-amu.fr

PASCALE HARINCK
CENTRE MATHÉMATIQUES LAURENT SCHWARTZ
CNRS-UMR 7640, UNIVERSITÉ PARIS-SACLAY
ÉCOLE POLYTECHNIQUE
91128 PALAISEAU
FRANCE
pascale.harinck@polytechnique.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.149

REGULARITY OF THE ANALYTIC TORSION FORM ON FAMILIES OF NORMAL COVERINGS

BING KWAN SO AND GUANGXIANG SU

We prove the smoothness of the L^2-analytic torsion form for fiber bundles with positive Novikov–Shubin invariant. We do so by generalizing the arguments of Azzali, Goette and Schick to an appropriate Sobolev space, and proving that the Novikov–Shubin invariant is also positive in the Sobolev setting, using an argument of Alvarez Lopez and Kordyukov.

1. Introduction

Let M be a closed Riemannian manifold and F be a flat vector bundle on M. Ray and Singer [1971] introduced the analytic torsion, which is the analytic analogue of the combinatorial torsion (see [Milnor 1966]). Let $Z \to M \xrightarrow{\pi} B$ be a fiber bundle with connected closed fibers $Z_x = \pi^{-1}(x)$ and F be a flat complex vector bundle on M with a flat connection ∇^F and a Hermitian metric h^F. Let T^HM be a horizontal distribution for the fiber bundle and g^{TZ} be a vertical Riemannian metric. Bismut and Lott [1995, (3.118)] introduced the torsion form $\mathcal{T}(T^HM, g^{TZ}, h^F) \in \Omega(B)$ defined by

$$\mathcal{T}(T^HM, g^{TZ}, h^F) = -\int_0^{+\infty} \Big[f^\wedge\big(C'_t, h^W\big) - \tfrac{1}{2}\chi'(Z;F) f'(0) \\ -\big(\tfrac{1}{4}\dim(Z)\operatorname{rk}(F)\chi(Z) - \tfrac{1}{2}\chi'(Z;F)\big) f'\big(\tfrac{1}{2} i\sqrt{t}\big)\Big] \frac{dt}{t}. \tag{1}$$

See [Bismut and Lott 1995] for the meaning of the terms in the integrand. To show the integral in the above formula is well defined, one must calculate the asymptotic of $f^\wedge(C'_t, h^W)$ as $t \to 0$ and the asymptotic as $t \to \infty$. For the asymptotic as $t \to 0$, they used the local index technique. For the asymptotic as $t \to \infty$, the key fact is that the fiber Z is closed, so the fiberwise operators involved have uniform positive lower bound for positive eigenvalues. They also proved a C^∞-analogue of the Riemann–Roch–Grothendieck theorem and proved that the torsion form is the transgression of the Riemann–Roch–Grothendieck theorem (see [Bismut and Lott 1995, Theorem 3.23]) and showed the zero degree part of $\mathcal{T}(T^HM, g^{TZ}, h^F) \in \Omega(B)$ is the Ray–Singer analytic torsion (see their Theorem 3.29).

MSC2010: primary 58J52; secondary 58J35.
Keywords: L^2-analytic torsion form, Novikov-Shubin invariant, Sobolev space, superconnection.

On the other hand, the L^2-analytic torsion was defined and studied by several authors; see [Carey and Mathai 1992], [Lott 1992], [Mathai 1992], etc. So it is natural to extend the L^2-analytic torsion to the family case, that is, to define and study the Bismut–Lott torsion form when the fiber Z is noncompact. From the above we see that one must study the asymptotic of the L^2 analogue of $f^\wedge\big(C_t, h^W\big)$ as $t \to 0$ and $t \to \infty$. Since in the L^2 case $f^\wedge\big(C'_t, h^W\big)$ has the same asymptotic as $t \to 0$, this part is easy. But in general the integral at ∞ does not converge, since in the L^2 case the positive eigenvalues of the fiberwise operator involved in $f^\wedge\big(C'_t, h^W\big)$ may not have a positive lower bound. To overcome this difficulty, one considers the special case where the Novikov–Shubin invariant is (sufficiently) positive. Gong and Rothenberg [1996] defined the L^2-analytic torsion form and proved that the torsion form is smooth, under the condition that the Novikov–Shubin invariant is at least half of the dimension of the base manifold. Heitsch and Lazarov [2002] generalized essentially the same arguments to foliations. In [Azzali et al. 2015], Azzali, Goette and Schick proved that the integrand defining the L^2-analytic torsion form, as well as several other invariants related to the signature operator, converges, provided the Novikov–Shubin invariant is positive (or of determinant class and L^2-acyclic). However, they did not prove the smoothness of the L^2-analytic torsion form. To consider the transgression formula, they had to use weak derivatives.

The aim of this paper is to establish the regularity of the L^2-analytic torsion form in the case when the Novikov–Shubin invariant is positive. Our motivation comes from the study of analytic torsion on some "noncommutative" spaces (along the lines of [Gorokhovsky and Lott 2006], etc., for local index). In this case one considers universal differential forms (as in the same paper), and Duhamel's formula for the heat operator having infinitely many terms. Instead, one makes essential use of the results of [Azzali et al. 2015] to ensure that (1) is well defined in the noncommutative case. We achieve this result by generalizing Azzali, Goette and Schick's arguments to some Sobolev spaces.

The rest of the paper is organized as follows. In Section 2, we define Sobolev norms on the spaces of kernels on the fibered product groupoid. Unlike [Azzali et al. 2015], we consider Hilbert–Schmidt type norms on the space of smoothing operators. Given a kernel, the Hilbert–Schmidt norm can be explicitly written down. As a result, we are able to take into account derivatives in both the fiberwise and transverse directions, with the help of a splitting similar to [Heitsch 1995]. In Section 3, we turn to proving that having positive Novikov–Shubin invariant implies positivity of the Novikov–Shubin invariant in the Sobolev settings. We adapt an argument of Alvarez Lopez and Kordyukov [2001]. In Section 4, we apply the arguments in [Azzali et al. 2015] and conclude that the integral in equation (1) converges in all Sobolev norms, and hence obtain the regularity of the L^2-analytic torsion form.

2. Preliminaries

In this section, we will define Sobolev norms on the space of kernels on the fibered product groupoid.

2A. ***The geometric setting.*** Let $Z \to M \xrightarrow{\pi} B$ be a fiber bundle with connected fibers $Z_x = \pi^{-1}(x)$, $x \in B$. Let $E \xrightarrow{\wp} M$ be a vector bundle. We assume B is compact. Let $V := \mathrm{Ker}(d\pi) \subset TM$.

We suppose that there is a finitely generated discrete group G acting on M from the right freely and properly discontinuously. We also assume that the group G acts on B such that the actions commute with π and $M_0 := M/G$ is a compact manifold. Since the submersion π is G-invariant, M_0 is also foliated and we denote its foliation V_0. Fix a distribution $H_0 \subset TM_0$ complementary to V_0. Fix a metric on V_0 and take a G-invariant metric on B, then these induce a Riemannian metric on M_0 by $g^{V_0} \oplus \pi^* g^{TB}$ on $TM_0 = V_0 \oplus H_0$.

Since the projection from M to M_0 is a local diffeomorphism, one gets a G-invariant splitting $TM = V \oplus H$. Denote by P^V and P^H respectively the projections to V and H. Moreover, one gets a G-invariant metric on V and a Riemannian metric on M by $g^{TM} = g^V \oplus \pi^* g^{TB}$ on $TM = V \oplus H$.

Given any vector field $X \in \Gamma^\infty(TB)$, denote the horizontal lift of X by $X^H \in \Gamma^\infty(H) \subset \Gamma^\infty(TM)$. By our construction, $|X^H|_{g_M}(p) = |X|_{g_B}(\pi(p))$.

Denote by μ_x and μ_B respectively the Riemannian measures on Z_x and B.

Definition 2.1. Let $E \xrightarrow{\wp} M$ be a complex vector bundle. We say that E is a contravariant G-bundle if G also acts on E from the right, such that for any $v \in E$, $g \in G$, $\wp(vg) = \wp(v)g \in M$, and moreover G acts as a linear map between the fibers.

The group G then acts on sections of E from the left by

$$s \mapsto g^* s, \quad (g^* s)(p) := s(pg)g^{-1} \in \wp^{-1}(p), \qquad \text{for all } p \in M.$$

We assume that E is endowed with a G-invariant metric g_E, and a G-invariant connection ∇^E (which is obviously possible if E is the pullback of some bundle on M_0). In particular, for any G-invariant section s of E, $|s|$ is a G-invariant function on M.

In the following, for any vector bundle F we denote its dual bundle by F'.

Recall that the "infinite dimensional bundle" over B in the sense of Bismut is a vector bundle with typical fiber $\Gamma_c^\infty(E|_{Z_x})$ (or other function spaces) over each $x \in B$. We denote by $E_\flat$ such a Bismut bundle. The space of smooth sections on $E_\flat$ is, as a vector space, $\Gamma_c^\infty(E)$. Each element $s \in \Gamma_c^\infty(E)$ is regarded as a map

$$x \mapsto s|_{Z_x} \in \Gamma_c^\infty\big(E|_{Z_x}\big) \quad \text{for all } x \in B.$$

In other words, one defines a section on $E_\flat$ to be smooth if the images of all $x \in B$ fit together to form an element in $\Gamma^\infty_c(E)$. In particular, $\Gamma^\infty_c((M \times \mathbb{C})_\flat) = C^\infty_c(M)$, and one identifies $\Gamma^\infty_c(TB \otimes (M \times \mathbb{C})_\flat)$ with $\Gamma^\infty_c(H)$ by $X \otimes f \mapsto f X^H$.

2B. ***Covariant derivatives and Sobolev spaces.*** Let ∇^E be a G-invariant connection on E. Denote by ∇^{TM}, ∇^{TB} the Levi-Civita connections (with respect to the metrics defined in the last section). Note that $[X^H, Y] \in \Gamma^\infty(V)$ for any vertical vector field $Y \in \Gamma^\infty(V)$. One naturally defines the connections

$$\begin{aligned} \nabla^{V_\flat}_X Y &:= [X^H, Y] \quad \text{for all } Y \in \Gamma^\infty(V_\flat) \cong \Gamma^\infty(V), \\ \nabla^{E_\flat}_X s &:= \nabla^E_{X^H} s \qquad \text{for all } s \in \Gamma^\infty(E_\flat) \cong \Gamma^\infty_c(E). \end{aligned}$$

Definition 2.2. The covariant derivative on $E_\flat$ is the map

$$\dot\nabla^{E_\flat} : \Gamma^\infty\Big(\otimes^\bullet T^*B \bigotimes \otimes^\bullet V'_\flat \bigotimes E_\flat\Big) \to \Gamma^\infty\Big(\otimes^{\bullet+1} T^*B \bigotimes \otimes^\bullet V'_\flat \bigotimes E_\flat\Big),$$

defined by

$$\begin{aligned} (2)\quad & \big(\dot\nabla^{E_\flat} s\big)(X_0, X_1, \dots, X_k; Y_1, \dots, Y_l) \\ &:= \nabla^{E_\flat}_{X_0} s(X_1, \dots, X_k; Y_1, \dots, Y_l) - \sum_{j=1}^{l} s\big(X_1, \dots, X_k; Y_1, \dots, \nabla^{V_\flat}_{X_0} Y_j, \dots, Y_l\big) \\ &\qquad - \sum_{i=1}^{k} s\big(X_1, \dots, \nabla^{TB}_{X_0} X_i, \dots, X_k; Y_1, \dots Y_l\big), \end{aligned}$$

for any $k, l \in \mathbb{N}$, $X_0, \dots, X_k \in \Gamma^\infty(TB)$, $Y_1, \dots, Y_l \in \Gamma^\infty(V)$.

Clearly, taking the covariant derivative can be iterated, which we denote by $(\dot\nabla^{E_\flat})^m$, $m = 1, 2, \dots$. Note that $(\dot\nabla^{E_\flat})^m$ is a differential operator of order m.

Also, we define

$$\dot\partial^V : \Gamma^\infty\Big(\otimes^\bullet T^*B \bigotimes \otimes^\bullet V'_\flat \bigotimes E_\flat\Big) \to \Gamma^\infty\Big(\otimes^\bullet T^*B \bigotimes \otimes^{\bullet+1} V'_\flat \bigotimes E_\flat\Big)$$

by

$$\begin{aligned} (3)\quad & \big(\dot\partial^V s\big)(X_1, \dots, X_k; Y_0, Y_1, \dots, Y_l) \\ &:= \nabla^E_{Y_0} s(X_1, \dots, X_k; Y_1, \dots, Y_l) - \sum_{j=1}^{l} s(X_1, \dots, X_k; Y_1, \dots, P^V(\nabla^{TM}_{Y_0} Y_j), \dots, Y_l). \end{aligned}$$

In the following definition, we regard $(\dot\nabla^{E_\flat})^i (\dot\partial^V)^j s \in \Gamma^\infty(\otimes^i H' \bigotimes \otimes^j V' \bigotimes E_\flat)$.

Definition 2.3. For $s \in \Gamma^\infty_c(E)$, we define its m-th Sobolev norm by

$$(4)\qquad \|s\|^2_m := \sum_{i+j \le m} \int_{x \in B} \int_{y \in Z_x} \big|(\dot\nabla^{E_\flat})^i (\dot\partial^V)^j s\big|^2 (x, y) \mu_x(y) \mu_B(x).$$

Denote by $\mathcal{W}^m(E)$ the Sobolev completion of $\Gamma_c^\infty(E)$ with respect to $\|\cdot\|_m$.

Recall that an operator A is called C^∞-bounded if in normal coordinates the coefficients and their derivatives are C^∞-bounded.

Since M is locally isometric to a compact space M_0, it is a manifold with bounded geometry (see [Shubin 1992, Appendix 1] for an introduction). Moreover, ∇^E is a C^∞-bounded differential operator, because by G-invariance the Christoffel symbols of ∇^E and all their derivatives are uniformly bounded. Using normal coordinate charts and parallel transport with respect to ∇^E as the trivialization, one sees that E is a bundle with bounded geometry.

Since the operators $\dot{\nabla}^{E_\flat}$ and $\dot{\partial}^V$ are just respectively the (0, 1) and (1, 0) parts of the usual covariant derivative operator, our Definition 2.3 is equivalent to the standard Sobolev norm [Shubin 1992, Appendix 1 (1.3)] (with $p = 2$ and s nonnegative integers).

One has elliptic regularity for these Sobolev spaces:

Lemma 2.4 [Shubin 1992, Lemma 1.4]. *Let A be any C^∞-bounded, uniformly elliptic differential operator of order m. For any $i, j \geq 0$, there exists a constant C such that for any $s \in \Gamma_c^\infty(E)$*

$$\|s\|_{i+m} \leq C(\|As\|_i + \|s\|_j).$$

Remark 2.5. Throughout this paper, by an "elliptic operator" on a manifold, we mean elliptic in all directions, without taking any foliation structure into consideration. We use the term "fiberwise elliptic operators" to refer to differential operators that are fiberwise and elliptic restricted to fibers.

2C. *The fibered product.*

Definition 2.6. The fibered product of the manifold M is

$$M \times_B M := \{(p, q) \in M \times M : \pi(p) = \pi(q)\},$$

and with the maps from $M \times_B M$ to M defined by

$$\boldsymbol{s}(p, q) := q \quad \text{and} \quad \boldsymbol{t}(p, q) := p.$$

The manifold $M \times_B M$ is a fiber bundle over B, with typical fiber $Z \times Z$. One naturally has the splitting [Heitsch 1995, Section 2]

$$T(M \times_B M) = \hat{H} \oplus V_{\boldsymbol{t}} \oplus V_{\boldsymbol{s}},$$

where $V_{\boldsymbol{s}} := \mathrm{Ker}(d\boldsymbol{t})$ and $V_{\boldsymbol{t}} := \mathrm{Ker}(d\boldsymbol{s})$.

Note that $V_{\boldsymbol{s}} \cong \boldsymbol{s}^* V$ and $V_{\boldsymbol{t}} \cong \boldsymbol{t}^* V$. As in Section 2A, we endow $M \times_B M$ with a metric by lifting the metrics on H_0 and V_0. Then $M \times_B M$ is a manifold with bounded geometry.

Notation 2.7. With some abuse in notation, we shall often write elements in $M \times_B M$ as a triple (x, y, z) and $\boldsymbol{s}(x, y, z) = (x, z)$, $\boldsymbol{t}(x, y, z) = (x, y) \in M$, where $x \in B$ and $y, z \in Z_x$.

Let G act on $M \times_B M$ by the diagonal action $(p, q)g := (pg, qg)$. Let $E \to M$ be a contravariant G-vector bundle and E' be its dual. We shall consider

$$\hat{E} \to M \times_B M := \boldsymbol{t}^* E \otimes \boldsymbol{s}^* E'.$$

Given a G-invariant connection ∇^E on E, let

$$\nabla^{\hat{E}} := \boldsymbol{t}^* \nabla^E \otimes \mathrm{id}_{\boldsymbol{s}^* E'} + \mathrm{id}_{\boldsymbol{t}^* E} \otimes \boldsymbol{s}^* \nabla^{E'}$$

be the tensor product of the pullback connections. Fix any local base $\{e_1, \dots e_r\}$ of E' on some $U \subset M$. Any section can be written as

$$s = \sum_{i=1}^{r} u_i \otimes \boldsymbol{s}^* e_i$$

on $\boldsymbol{s}^{-1}(U)$, where $u_i \in \Gamma^\infty(\boldsymbol{t}^* E)$. Then by definition we have for any vector X on M,

$$\nabla^{\hat{E}}_X \left(\sum_{i=1}^{r} u_i \otimes \boldsymbol{s}^* e_i \right) = \sum_{i=1}^{r} (\nabla^{\boldsymbol{t}^* E}_X u_i) \otimes \boldsymbol{s}^* e_i + u_i \otimes \boldsymbol{s}^* (\nabla^{E'}_{ds(X)} e_i). \tag{5}$$

Similarly to Definition 2.2, we define the covariant derivative operators on $\Gamma^\infty(\otimes^\bullet T^* B \bigotimes \otimes^\bullet (V'_{\boldsymbol{t}})_\flat \bigotimes \otimes^\bullet (V'_{\boldsymbol{s}})_\flat \bigotimes \hat{E}_\flat)$.

Definition 2.8. Define

$$\begin{aligned}
&\big(\dot{\nabla}^{\hat{E}_\flat} \psi\big)(X_0, X_1, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}) \\
&\qquad := \nabla^{\hat{E}_\flat}_{X_0} \psi(X_1, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}) \\
&\qquad\qquad - \sum_{1 \le j \le l} \psi(X_1, \dots, X_k; Y_1, \dots, \nabla^{V_\flat}_{X_0} Y_j, \dots, Y_l, Z_1, \dots, Z_{l'}) \\
&\qquad\qquad - \sum_{1 \le j \le l'} \psi(X_1, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, \nabla^{V_\flat}_{X_0} Z_j, \dots, Z_{l'}) \\
&\qquad\qquad - \sum_{1 \le i \le k} \psi(X_1, \dots, \nabla^{TB}_{X_0} X_i, \dots, X_k; Y_1, \dots Y_l, Z_1, \dots, Z_{l'}),
\end{aligned}$$

and

$$
\begin{aligned}
\big(\dot{\partial}^s\psi\big)&(X_1,\dots,X_k;Y_0,Y_1,\dots,Y_l,Z_1,\dots,Z_{l'})\\
&:=\nabla^{\hat{E}}_{Y_0}\psi(X_1,\dots,X_k;Y_1,\dots,Y_l,Z_1,\dots,Z_{l'})\\
&\qquad-\sum_{1\le j\le l}\psi(X_1,\dots,X_k;Y_1,\dots,P^{V^s}(\nabla^{TM}_{Y_0}Y_j),\dots,Y_l,Z_1,\dots,Z_{l'})\\
&\qquad-\sum_{1\le j\le l'}\psi(X_1,\dots,X_k;Y_1,\dots,Y_l,Z_1,\dots,P^{V^t}[Y_0,Z_j],\dots,Z_{l'}),
\end{aligned}
$$

and

$$
\begin{aligned}
\big(\dot{\partial}^t\psi\big)&(X_1,\dots,X_k;Y_1,\dots,Y_l,Z_0,Z_1,\dots,Z_{l'})\\
&:=\nabla^{\hat{E}}_{Y_0}\psi(X_1,\dots,X_k;Y_1,\dots,Y_l,Z_0,Z_1,\dots,Z_{l'})\\
&\qquad-\sum_{1\le j\le l}\psi(X_1,\dots,X_k;Y_1,\dots,P^{V^s}[Z_0,Y_j],\dots,Y_l,Z_1,\dots,Z_{l'})\\
&\qquad-\sum_{1\le j\le l'}\psi(X_1,\dots,X_k;Y_1,\dots,Y_l,Z_1,\dots,P^{V^t}(\nabla^{TM}_{Z_0}Z_j),\dots,Z_{l'}).
\end{aligned}
$$

Given any vector fields Y, $Z\in V$, let Y^s, Z^t be respectively the lifts of Y and Z to V_s and V_t. Then $[Y^s, Z^t]=0$. It follows that as differential operators, $[\dot{\partial}^s, \dot{\partial}^t]=0$. Also, it is straightforward to verify that $[\dot{\nabla}^{\hat{E}_\flat}, \dot{\partial}^s]$ and $[\dot{\nabla}^{\hat{E}_\flat}, \dot{\partial}^t]$ are both zeroth order differential operators (i.e., smooth bundle maps).

Fix a local trivialization

$$\boldsymbol{x}_\alpha:\pi^{-1}(B_\alpha)\to B_\alpha\times Z,\quad p\mapsto(\pi(p),\varphi^\alpha(p)),$$

where $B=\bigcup_\alpha B_\alpha$ is a finite open cover (since B is compact), and $\varphi^\alpha|_{\pi^{-1}(x)}:Z_x\to Z$ is a diffeomorphism. Such a trivialization induces a local trivialization of the fiber bundle $M\times_B M\xrightarrow{t} M$ by $M=\bigcup M_\alpha$, $M_\alpha:=\pi^{-1}(B_\alpha)$,

$$\hat{\boldsymbol{x}}_\alpha:\boldsymbol{t}^{-1}(M_\alpha)\to M_\alpha\times Z,\quad (p,q)\mapsto(p,\varphi^\alpha(q)).$$

On $M_\alpha\times Z$ the source and target maps are explicitly given by

$$\boldsymbol{s}\circ(\hat{\boldsymbol{x}}_\alpha)^{-1}(p,z)=(\boldsymbol{x}_\alpha)^{-1}(\pi(p),z)\text{ and }\boldsymbol{t}\circ(\hat{\boldsymbol{x}}_\alpha)^{-1}(p,z)=p.\tag{6}$$

For such a trivialization, one has the natural splitting

$$T(M_\alpha\times Z)=H^\alpha\oplus V^\alpha\oplus TZ,$$

where H^α and V^α are respectively H and V restricted to $M_\alpha\times\{z\}$, $z\in Z$. It follows from (6) that

$$V^\alpha=d\hat{\boldsymbol{x}}_\alpha(V_s)\quad\text{and}\quad TZ=d\hat{\boldsymbol{x}}_\alpha(V_t).$$

Given any vector field X on B, let X^H, $X^{\hat{H}}$ be respectively the lifts of X to H and $\hat{H}$. Since $d\boldsymbol{t}(X^{\hat{H}}) = d\boldsymbol{s}(X^{\hat{H}}) = X^H$, it follows that

$$d\hat{\boldsymbol{x}}_\alpha(X^{\hat{H}}) = X^{H^\alpha} + d\varphi^\alpha(X^H).$$

Note that $d\varphi^\alpha(X^H) \in TZ \subseteq T(M_\alpha \times Z)$.

Corresponding to the splitting $T(M_\alpha \times Z) = H^\alpha \oplus V^\alpha \oplus TZ$, one can define the covariant derivative operators. Let ∇^{TM_α} be the Levi-Civita connection on M_α and ∇^{TZ} be the Levi-Civita connection on Z. Define for any smooth section $\phi \in \Gamma^\infty(\otimes^\bullet T^*B \bigotimes \otimes^\bullet (V^\alpha)'_\flat \bigotimes \otimes^\bullet T^*Z_\flat \bigotimes (\hat{\boldsymbol{x}}_\alpha^{-1})^*\hat{E}_\flat)$,

$$\begin{aligned}
\big(\dot{\nabla}^\alpha \phi\big)&(X_0, X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&:= (\boldsymbol{x}_\alpha^* \nabla^{\hat{E}_\flat})_{X_0^{H^\alpha}} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l} \phi(X_1, \ldots, X_k; Y_1, \ldots, [X_0^{H^\alpha}, Y_j], \ldots, Y_l, , Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l'} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, [X_0^{H^\alpha} Z_j], \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le i \le k} \phi(X_1, \ldots, \nabla^{TB}_{X_0} X_i, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, Z_{l'}),
\end{aligned}$$

and

$$\begin{aligned}
\big(\dot{\partial}^\alpha \phi\big)&(X_1, \ldots, X_k; Y_0, Y_1, \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&:= (\boldsymbol{x}_\alpha^* \nabla^{\hat{E}_\flat})_{Y_0} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l} \phi(X_1, \ldots, X_k; Y_1, \ldots, P^{V^\alpha}(\nabla^{TM_\alpha}_{Y_0} Y_j), \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l'} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, P^{TZ}[Y_0, Z_j], \ldots, Z_{l'}),
\end{aligned}$$

and

$$\begin{aligned}
\big(\dot{\partial}^Z \phi\big)&(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_0, Z_1, \ldots, Z_{l'}) \\
&:= (\boldsymbol{x}_\alpha^* \nabla^{\hat{E}_\flat})_{Z_0} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_0, Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l} \phi(X_1, \ldots, X_k; Y_1, \ldots, P^{V^\alpha}[Z_0, Y_j], \ldots, Y_l, Z_1, \ldots, Z_{l'}) \\
&\quad - \sum_{1 \le j \le l'} \phi(X_1, \ldots, X_k; Y_1, \ldots, Y_l, Z_1, \ldots, \nabla^{TZ}_{Z_0} Z_j, \ldots, Z_{l'}).
\end{aligned}$$

Consider the special case when $\phi = u \otimes \boldsymbol{s}^* e$, where $u \in \Gamma^\infty(\otimes^\bullet T^*B \bigotimes \otimes^\bullet (V^\alpha)'_\flat \otimes \boldsymbol{t}^* E)$ and $e \in \Gamma^\infty(\otimes^\bullet T^* Z_\flat \otimes E')$.

Lemma 2.9. *For* $(x, y, z) \in M_\alpha \times Z$, *one has*

$$\dot{\nabla}^\alpha(u \otimes \boldsymbol{s}^* e)(x, y, z) = \big(\dot{\nabla}^{E_\flat} u|_{M_\alpha \times \{z\}}(x, y)\big) \otimes \boldsymbol{s}^*(e(x, z)) + u \otimes \boldsymbol{s}^*\big(\nabla^{E'_\flat} e(x, z)\big)$$

and $\dot{\partial}^\alpha(u \otimes \boldsymbol{s}^* e)(x, y, z) = (\dot{\partial}^V u|_{M_\alpha \times \{z\}}(x, y)) \otimes \boldsymbol{s}^*(e(x, z))$.

Proof. It suffices to consider the case when Y_j, $Z_{j'}$ are respectively vector fields on M_α and Z lifted to $M_\alpha \times Z$. From this assumption it follows that $[Y_j, Z_{j'}] = [X_0^{H^\alpha}, Z_{j'}] = 0$. The lemma follows by a simple computation. $\square$

We express the (pullback of) covariant derivatives $\dot{\nabla}^{\hat{E}_\flat}\psi$, $\dot{\partial}^{\boldsymbol{s}}\psi$ and $\dot{\partial}^{\boldsymbol{t}}\psi$ in terms of $\dot{\nabla}^\alpha \psi^\alpha$, $\dot{\partial}^\alpha \psi^\alpha$ and $\dot{\partial}^Z \psi^\alpha$, where $\psi^\alpha := (\boldsymbol{x}_\alpha^{-1})^* \psi$. One directly verifies

$$
\begin{aligned}
(7)\quad & \big(\dot{\nabla}^{E_\flat}\psi\big)(X_0, X_1, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}) \\
&= (\boldsymbol{x}_\alpha^{-1})^* \Bigg(\nabla^\alpha_{(X_0^{H^\alpha} + d\varphi^\alpha(X_0^H))} \psi^\alpha\big(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_1, \dots, Y_l, Z_1, \dots, Z_{l'})\big) \\
&\qquad - \sum_{1 \le j \le l} \psi^\alpha\big(X_1, \dots, X_k; d\boldsymbol{x}_\alpha Y_1, \dots, [X_0^{H^\alpha}, d\boldsymbol{x}_\alpha Y_j], \dots, d\boldsymbol{x}_\alpha Y_l, d\boldsymbol{x}_\alpha(Z_1, \dots, Z_{l'})\big) \\
&\quad - \sum_{1 \le j \le l'} \psi^\alpha\big(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_1, \dots, Y_l), d\boldsymbol{x}_\alpha Z_1, \dots, [X_0^{H^\alpha} + d\varphi^\alpha(X_0^H), d\boldsymbol{x}_\alpha Z_j], \dots, d\boldsymbol{x}_\alpha Z_{l'}\big) \\
&\qquad\qquad - \sum_{1 \le i \le k} \psi^\alpha\big(X_1, \dots, \nabla^{TB}_{X_0} X_i, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}\big) \Bigg) \\
&= (\boldsymbol{x}_\alpha^{-1})^* \big(\dot{\nabla}^\alpha \psi^\alpha(X_0, X_1, \dots, X_k; Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}) \\
&\qquad\qquad + \dot{\partial}^Z \psi^\alpha(X_1, \dots, X_k; Y_1, \dots, Y_l, d\varphi^\alpha(X_0^H), Z_1, \dots, Z_{l'}) \\
&\quad + \sum_{1 \le j \le l'} \psi^\alpha\big(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_1, \dots, Y_l), d\boldsymbol{x}_\alpha Z_1, \dots, (\nabla^{TZ} d\varphi^\alpha(X_0^H))(d\boldsymbol{x}_\alpha Z_j), \dots, d\boldsymbol{x}_\alpha Z_{l'}\big)\big).
\end{aligned}
$$

By similar computations for $\dot{\partial}^{\boldsymbol{s}}$ and $\dot{\partial}^{\boldsymbol{t}}$, one gets:

$$
\begin{aligned}
(8)\quad & \big(\dot{\partial}^{\boldsymbol{s}}\psi\big)(X_1, \dots, X_k; Y_0, Y_1, \dots Y_l, Z_1, \dots Z_{l'}) \\
&\qquad = (\boldsymbol{x}_\alpha^{-1})^* \big(\dot{\partial}^\alpha \psi^\alpha(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_0, Y_1, \dots, Y_l, Z_1, \dots, Z_{l'}) \big),
\end{aligned}
$$

and

$$
\begin{aligned}
(9)\quad & \big(\dot{\partial}^{\boldsymbol{t}}\psi\big)(X_1, \dots, X_k; Y_1, \dots, Y_l, Z_0, Z_1, \dots, Z_{l'}) \\
&\qquad = (\boldsymbol{x}_\alpha^{-1})^* \big(\dot{\partial}^Z \psi^\alpha(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_1, \dots, Y_l, Z_0, \dots, Z_{l'})) \big) \\
&\qquad\quad + \sum_{1 \le j \le l'} \psi^\alpha\big(X_1, \dots, X_k; d\boldsymbol{x}_\alpha(Y_1, \dots, Y_l), d\boldsymbol{x}_\alpha Z_1, \\
&\qquad\qquad \dots, \big(\nabla^{TZ}_{d\boldsymbol{x}_\alpha Z_0} d\boldsymbol{x}_\alpha Z_j - d\boldsymbol{x}_\alpha(P^{V_t} \nabla^{TM}_{Z_0} Z_j), \dots, d\boldsymbol{x}_\alpha Z_{l'}\big)\big).
\end{aligned}
$$

2D. ***Smoothing operators.*** For any $(x, y, z) \in M \times_B M$, let $\boldsymbol{d}(x, y, z)$ be the Riemannian distance between $y, z \in Z_x$. We regard $\boldsymbol{d}$ as a continuous, nonnegative function on $M \times_B M$.

Definition 2.10. (See [Nistor et al. 1999]). As a vector space,

$$\Psi^{-\infty}_{\infty}(M \times_B M, E) := \left\{ \psi \in \Gamma^{\infty}(\hat{E}) : \begin{array}{l} \text{For any } m \in \mathbb{N}, \varepsilon > 0, \text{there exists } C_m > 0 \\ \text{such that for all } i + j + k \leq m, \\ |(\dot{\nabla}^{\hat{E}_\flat})^i (\dot{\partial}^{V_s})^j (\dot{\partial}^{V_t})^k \psi| \leq C_m e^{-\varepsilon \boldsymbol{d}}. \end{array} \right\}.$$

The convolution product structure on $\Psi^{-\infty}_{\infty}(M \times_B M, E)$ is defined by

$$\psi_1 \star \psi_2(x, y, z) := \int_{Z_x} \psi_1(x, y, w)\psi_2(x, w, z)\mu_x(w).$$

We introduce a Sobolev type generalization of the Hilbert–Schmidt norm on $\Psi^{-\infty}_{\infty}(M \times_B M, E)^G$, the space of G-invariant kernels. Since G is a finitely generated discrete group and acts on M freely and properly discontinuously, then there exists a smooth compactly supported function $\chi \in C^\infty_c(M)$, such that

$$\sum_{g \in G} g^* \chi = 1.$$

In particular, one may construct χ as follows. Denote by π_G the projection $M \to M_0 = M/G$. There exists some $r > 0$ and a finite collection of geodesic balls $B(p_\alpha, r)$ of radius r such that $B(p_\alpha, r)$ is diffeomorphic to its image in M_0 under π_G, and moreover $\{B(p_\alpha, \frac{1}{3}r)\}$ covers M_0 (since M_0 is compact). Since G acts on M by isometry, $\pi_G(B(p_\alpha g, r)) = \pi_G(B(p_\alpha, r))$ for all $g \in G$. Thus one may without loss of generality assume that $B(p_\alpha, r)$ are mutually disjoint.

Define the functions $f \in C^\infty(\mathbb{R})$, $F_\alpha, F \in C^\infty_c(M)$ by

$$\begin{aligned} f(t) &:= e^{-1/t^2} \text{ if } t > 0, \quad 0 \text{ if } t \leq 0, \\ F_\alpha(p) &:= f\Big(1 - \frac{2}{r}\boldsymbol{d}(p, p_\alpha)\Big)\Big(f\Big(\frac{3}{r}\boldsymbol{d}(p, p_\alpha) - 1\Big) + f\Big(1 - \frac{2}{r}\boldsymbol{d}(p, p_\alpha)\Big)\Big)^{-1}, \ p \in M, \\ F &:= \sum_\alpha F_\alpha. \end{aligned}$$

Note that F is well defined because F_α is supported on $B(p_\alpha, r)$, which is locally finite. Since by construction

$$\Big\{\bigcup_\alpha B\Big(p_\alpha g, \frac{r}{3}\Big)\Big\}_{g \in G}$$

is a locally finite cover of M, $\sum_g g^*F$ is also well defined. Define

$$\chi := F\Big(\sum_g g^*F\Big)^{-1}.$$

Then clearly χ is the required partition of unity. Moreover, observe that $\chi^{1/2}$ is a smooth function because $f^{1/2}$ is smooth and all denominators are uniformly bounded away from 0.

For any G-invariant $\psi \in \Psi^{-\infty}_{\infty}(M \times_B M, E)^G$, recall that the standard trace of ψ is

$$\mathrm{tr}_\Psi(\psi)(x) := \int_{z \in Z_x} \chi(x,z)\,\mathrm{tr}(\psi(x,z,z))\mu_x(z) \in C^\infty(B).$$

The definition does not depend on the choice of χ. The corresponding Hilbert–Schmidt norm is

$$\begin{aligned}(10)\quad &\int_B \big(\mathrm{tr}_\Psi(\psi\psi^*)(x)\big)^2 \mu_B(x)\\ &\qquad = \int_B \int_{Z_x} \chi(x,z) \int_{Z_x} \mathrm{tr}(\psi(x,z,y)\psi^*(x,y,z))\mu_x(y)\mu_x(z)\mu_B(x).\end{aligned}$$

Note that equation (10) coincides with the L^2-norm of ψ. Generalizing (10) to taking into account derivatives, we define:

Definition 2.11. The m-th Hilbert–Schmidt norm on $\Psi^{-\infty}_{\infty}(M \times_B M, E)^G$ is defined to be

$$\|\psi\|^2_{\mathrm{HS}m} := \sum_{i+j+k \le m} \int_B \int_{Z_x} \chi(x,z) \int_{Z_x} \big|(\dot{\nabla}^{\hat{E}_\flat})^i (\dot{\partial}^s)^j (\dot{\partial}^t)^k \psi\big|^2 (x,y,z)\mu_x(y)\mu_x(z)\mu_B(x),$$

for any G-invariant element ψ. Let $\bar{\Psi}^{-\infty}_m(M \times_B M, E)^G$ be the completion of $\Psi^{-\infty}_{\infty}(M \times_B M, E)^G$ with respect to $\|\cdot\|_{\mathrm{HS}m}$.

Similar to Lemma 2.4, one has elliptic regularity for the Hilbert–Schmidt norm:

Lemma 2.12. *Let A be a G-invariant, first order elliptic differential operator, then for any $m = 0, 1, \ldots,$ there exists a constant $C > 0$ such that*

$$\|\psi\|_{\mathrm{HS}m+1} \le C(\|A\psi\|_{\mathrm{HS}m} + \|\psi\|_m),$$

for all $\psi \in \Psi^{-\infty}_{\infty}(M \times_B M, E)^G$.

Proof. Define

$$S := \{g \in G : \chi(g^*\chi) \ne 0\}.$$

Then S is finite because $\{g^*\chi\}$ is a locally finite partition of unity.

Consider $(\chi(x,z))^{1/2}\psi$. By the Leibniz rule, one has

$$(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\chi^{1/2}\psi=\chi^{1/2}(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\psi$$

modulo terms involving lower derivatives in ψ. Since $(\chi(x,z))^{1/2}$ is smooth with bounded derivatives, there exists some $C_1>0$ such that for any $(x,y,z)\in M\times_B M$,

$$\left|\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\chi^{1/2}\psi\right|^2-\chi\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\psi\right|^2\right|(x,y,z)\\ \le\sum_{g\in S}g^*\chi\biggl(C_1\sum_{i+j+k\le m-1}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\psi\right|^2\biggr)(x,y,z). \tag{11}$$

Similarly, since $A\chi^{1/2}-\chi^{1/2}A$ is a C^∞-bounded tensor, one has

$$\left|\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k(A\chi^{1/2}\psi)\right|^2-\chi\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^kA\psi\right|^2\right|\\ \le\sum_{g\in S}g^*\chi\biggl(C_2\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\psi\right|^2\biggr). \tag{12}$$

Since the integrand is G-invariant, for any $g\in G$

$$\int_{M\times_B M}g^*\chi\sum_{i+j+k\le m-1}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k\psi\right|^2\mu_x(y)\mu_x(z)\mu_B(x)=\|\psi\|^2_{\mathrm{HS}\,m-1}.$$

Observe that A being G-invariant implies A is uniformly elliptic and C^∞-bounded. Therefore applying Lemma 2.4 for $(\chi(x,z))^{1/2}\psi$, there exists a constant $C_3>0$ such that

$$\int_{M\times_B M}\sum_{i+j+k\le m+1}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k(\chi^{1/2}\psi)\right|^2\mu_x(y)\mu_x(z)\mu_B(x)\\ \le C_3\biggl(\int_{M\times_B M}\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k(A\chi^{1/2}\psi)\right|^2\mu_x(y)\mu_x(z)\mu_B(x)\\ +\int_{M\times_B M}\sum_{i+j+k\le m}\left|(\dot\nabla^{\hat E_\flat})^i(\dot\partial^{\boldsymbol s})^j(\dot\partial^{\boldsymbol t})^k(\chi^{1/2}\psi)\right|^2\mu_x(y)\mu_x(z)\mu_B(x)\biggr).$$

Then by equations (11) and (12), we get the lemma. □

2E. ***Fiberwise operators.*** We turn to considering another class of operators and a different norm.

Definition 2.13. A fiberwise operator is a linear operator $A : \Gamma_c^\infty(E_\flat) \to \mathcal{W}^0(E)$ such that for all $x \in B$, and any sections $s_1, s_2 \in \Gamma_c^\infty(E_\flat)$,

$$(As_1)(x) = (As_2)(x),$$

whenever $s_1(x) = s_2(x)$.

We say that A is smooth if $A(\Gamma_c^\infty(E)) \subseteq \Gamma^\infty(E)$. A smooth fiberwise operator A is said to be bounded of order m if A extends to a bounded map from $\mathcal{W}^m(E)$ to itself.

Denote by $\|A\|_{\mathrm{op}m}$ the operator norm of $A : \mathcal{W}^m(E) \to \mathcal{W}^m(E)$.

Example 2.14. Examples of smooth fiberwise operators are $\Psi_\infty^{-\infty}(M \times_B M, E)$, acting on $\mathcal{W}^m(E)$ by vector representation, i.e.,

$$(\Psi s)(x, y) := \int_{Z_x} \psi(x, y, z) s(x, z) \mu_x(z).$$

Notation 2.15. For the fiberwise operator $A : \Gamma_c^\infty(E_\flat) \to \mathcal{W}^0(E)$ which is of the form given by Example 2.14, we denote its kernel by $A(x, y, z)$. We shall write

$$\|A\|_{\mathrm{HS}m} := \|A(x, y, z)\|_{\mathrm{HS}m},$$

provided $A(x, y, z) \in \bar{\Psi}_m^{-\infty}(M \times_B M, E)$.

The following lemma enables one to construct more fiberwise operators:

Lemma 2.16. *Let A be any first order, C^∞-bounded differential operator on M and $\Psi \in \Psi_\infty^{-\infty}(M \times_B M, E)$ be as in Example 2.14. Then $[A, \Psi]$ is a fiberwise operator in $\Psi_\infty^{-\infty}(M \times_B M, E)$.*

Proof. Since multiplication by a tensor or differentiation along V is fiberwise, all that remains is to consider operators of the form $\nabla^E_{X^H}$, for some vector field X on B. Let $L^{\nabla^E}_{X^H} = d^{\nabla^E} i_{X^H} + i_{X^H} d^{\nabla^E}$, where d^{∇^E} is the twisted de Rham operator. In the remainder of this paper, the Lie derivatives are all defined in this way.

Let $s \in \Gamma_c^\infty(E)$ be arbitrary. We first suppose that Z is orientable and μ_x is a volume form. By the decay condition in Definition 2.10, one can differentiate under the integral sign to get

$$\begin{aligned} A\Psi s(x, z) &= \int_{Z_x} L^{\nabla^{\hat{E}}}_{X^{\hat{H}}} (\psi(x, y, z) s(x, y) \mu_x(y)) \\ &= \int_{Z_x} \big(L^{\nabla^{\hat{E}}}_{X^{\hat{H}}} \psi(x, y, z)\big) s(x, y) \mu_x(y) + \int_{Z_x} \psi(x, y, z) \big(L^{\nabla^E}_{X^H} s(x, y)\big) \mu_x(y) \\ &\qquad + \int_{Z_x} \psi(x, y, z) s(x, y) \big(L^{\nabla^E}_{X^H} \mu_x(y)\big). \end{aligned}$$

The second term in the last line is just ΨAs. Hence the result.

For the general case, one can take a suitable partition of unity and integrate over local volume forms. Then one obtains a similar equation. □

Let A be a smooth fiberwise operator on $\Gamma_c^\infty(E_\flat)$. Then A induces a fiberwise operator $\hat{A}$ on $\Gamma_c^\infty(\hat{E}_\flat)$ by

$$\hat{A}(u \otimes \boldsymbol{s}^* e) := A(u|_{M_\alpha \times \{z\}}) \otimes (\boldsymbol{s}^* e) \tag{13}$$

on $\boldsymbol{t}^{-1}(M_\alpha) \cong M_\alpha \times Z$, for any sections $e \in \Gamma^\infty(E')$ and $u \in \Gamma^\infty(\boldsymbol{t}^* E)$, and $\psi = u \otimes \boldsymbol{s}^* e \in \Gamma_c^\infty(\hat{E})$.

Note that $\hat{A}$ is independent of trivialization since A is fiberwise, and for any α, β and $z \in Z$, the transition function $\boldsymbol{x}_\beta \circ (\boldsymbol{x}_\alpha)^{-1}$ maps the submanifold $Z_x \times \{z\}$ to $Z_x \times \{\varphi_x^\beta \circ (\varphi_x^\alpha)^{-1}(z)\}$ as the identity diffeomorphism. If Ψ is a kernel and A is a fiberwise smooth operator, $A\Psi$ is also a kernel and is given by $\hat{A}\Psi$.

2F. *The main theorem.* Suppose that A is smooth and bounded of order m for all $m \in \mathbb{N}$. Consider the covariant derivatives of $\hat{A}\psi$.

Theorem 2.17. *For any smooth bounded G-invariant operator A, there exist constants $C'_{1,1}$, $C'_{0,0} > 0$ such that for any $\psi \in \Psi_\infty^{-\infty}(M \times_B M)^G$ one has $\hat{A}\psi \in \Psi_\infty^{-\infty}(M \times_B M)^G$ and*

$$\|\hat{A}\psi\|_{\mathrm{HS}\,1} \le (C'_{1,1}\|A\|_{\mathrm{op}\,1} + C'_{1,0}\|A\|_{\mathrm{op}\,0})\|\psi\|_{\mathrm{HS}\,1}.$$

Proof. Fix a partition of unity $\{\theta_\alpha\} \in C_c^\infty(B)$ subordinate to $\{B_\alpha\}$. We still denote by $\{\theta_\alpha\}$ its pullback to M and $M \times_B M$. Fix any Riemannian metric on Z and denote the corresponding Riemannian measure by μ_Z. Then one writes

$$(\hat{\boldsymbol{x}}_\alpha)_\star(\mu_x \mu_B) = J_\alpha \mu_B \mu_Z,$$

for some smooth positive function J_α. Moreover, over any compact subset of $B_\alpha \times Z$, $1/J_\alpha$ is bounded.

Given any $\psi \in \Psi_\infty^{-\infty}(M \times_B M)^G$, let $\psi^\alpha := \hat{\boldsymbol{x}}_\alpha^*(\psi)$. The theorem clearly follows from the inequalities

$$\int_{B_\alpha}\int_{Z_x} \chi(x,z) \int_{Z_x} |\dot{\nabla}^\alpha \hat{A}(\theta_\alpha \psi^\alpha)|^2 \mu_x(y)\mu_x(z)\mu_B(x) \le (C_1\|A\|^2_{\mathrm{op}\,1} + C_2\|A\|^2_{\mathrm{op}\,0})\|\psi\|^2_{\mathrm{HS}\,1}, \tag{14}$$

$$\int_{B_\alpha}\int_{Z_x} \chi(x,z) \int_{Z_x} |\dot{\partial}^\alpha \hat{A}(\theta_\alpha \psi^\alpha)|^2 \mu_x(y)\mu_x(z)\mu_B(x) \le (C_1\|A\|^2_{\mathrm{op}\,1} + C_2\|A\|^2_{\mathrm{op}\,0})\|\psi\|^2_{\mathrm{HS}\,1}, \tag{15}$$

$$\int_{B}\int_{Z_x} \chi(x,z) \int_{y \in Z_x} |\dot{\partial}^Z \hat{A}(\theta_\alpha \psi^\alpha)|^2 \mu_x(y)\mu_x(z)\mu_B(x) \le \|A\|^2_{\mathrm{op}\,0}\|\psi\|^2_{\mathrm{HS}\,1}. \tag{16}$$

Let $Z = \bigcup_\lambda Z_\lambda$ be a locally finite cover. Then the support of $\chi\theta_\alpha$ lies in some finite subcover. Let χ_α be the characteristic function

$$\chi_\alpha(x, z) = \begin{cases} 1 & \text{if } (\chi\theta_\alpha)(x, z) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Without loss of generality we may assume $E'|_{Z_\lambda}$ are all trivial. For each λ fix an orthonormal basis $\{e_r^\lambda\}$ of $E'|_{B_\alpha \times Z_\lambda}$, and write $\psi^\alpha := \sum_r u_r^\lambda \otimes \boldsymbol{s}^* e_r^\lambda$. Using Lemma 2.9, one estimates the integrand of the left-hand side of equation (14). Then there exits a constant $C_3 > 0$ such that

$$\begin{aligned}
&\big|\dot\nabla^\alpha(\hat A\theta_\alpha\psi^\alpha)\big|^2(x, y, z) \\
&\qquad = \Big|\sum_r (\dot\nabla^{E_\flat} A\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)) \otimes \boldsymbol{s}^* e_r^\lambda + (A\theta_\alpha u_r^\lambda) \otimes \boldsymbol{s}^*(\nabla^E e_r^\lambda)\Big|^2 \\
&\qquad \le C_3 \sum_r \Big(\big|\dot\nabla^{E_\flat} A\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)\big|^2 + \big|(A\theta_\alpha u_r^\lambda) \otimes \boldsymbol{s}^*(\nabla^E e_r^\lambda)\big|^2\Big).
\end{aligned}$$

By integrating, one gets for some constants C_q, $q = 4, \dots, 10$, that

$$\begin{aligned}
&\int_{B_\alpha}\int_{Z_x} \chi(x, z) \int_{Z_x} |\dot\nabla^\alpha \hat A(\theta_\alpha\psi^\alpha)|^2 \mu_x(y)\mu_x(z)\mu_B(x) \\
&\quad \le C_4 \sum_\lambda \int_{Z_\lambda}\int_{B_\alpha}\int_{Z_x} \sum_r \big(\big|\dot\nabla^{E_\flat} A\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)\big|^2 \\
&\qquad\qquad + \big|(A\theta_\alpha u_r^\lambda) \otimes \boldsymbol{s}^*(\nabla^E e_r^\lambda)\big|^2\big)\mu_x(y)\mu_B(x)\mu_Z(z) \\
&\quad \le \sum_\lambda \int_{Z_\lambda}\int_{B_\alpha}\int_{Z_x} \sum_r \big(C_5\|A\|^2_{\mathrm{op}\,1}\big(\big|\dot\nabla^{E_\flat}\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)\big|^2 \\
&\qquad\qquad\qquad \big|\dot\partial^V\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)\big|^2 + \big|\theta_\alpha(u_r^\lambda|_{M_\alpha\times\{z\}})(x, y)\big|^2\big) \\
&\qquad\qquad + C_6\|A\|_{\mathrm{op}\,0}\big|\theta_\alpha u_r^\lambda\big|^2\big)\mu_x(y)\mu_B(x)\mu_Z(z) \\
&\quad \le \sum_\lambda \int_{Z_\lambda}\int_{B_\alpha}\int_{Z_x} J_\alpha(C_7\|A\|^2_{\mathrm{op}\,1} + C_8\|A\|_{\mathrm{op}\,0})\big(\big|\dot\nabla^\alpha\theta_\alpha\psi_\alpha\big|^2 \\
&\qquad\qquad + \big|\dot\partial^\alpha\theta_\alpha\psi_\alpha\big|^2 + \big|\dot\partial^Z\theta_\alpha\psi_\alpha\big|^2 + \big|\theta_\alpha\psi_\alpha\big|^2\big)\mu_x(y)\mu_B(x)\mu_Z(z) \\
&\quad \le \int_B\int_{Z_x} \chi_\alpha \int_{Z_x} \big(C_9\|A\|^2_{\mathrm{op}\,1} + C_{10}\|A\|_{\mathrm{op}\,0}\big)\big(\big|\dot\nabla^{\hat E_\flat}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2 \\
&\qquad\qquad + \big|\dot\partial^{\boldsymbol{s}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2 + \big|\dot\partial^{\boldsymbol{t}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2 + \big|\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2\big)\mu_x(y)\mu_x(z)\mu_B(x).
\end{aligned}$$

Now we use an argument similar to the proof of Lemma 2.12. Namely, write the integrand as a sum

$$\chi_\alpha\big(\big|\dot{\nabla}^{\hat{E}_\flat}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{s}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{t}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2\big)$$
$$=\sum_{g\in S}\chi_\alpha g^*\chi\big(\big|\dot{\nabla}^{\hat{E}_\flat}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{s}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{t}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2\big).$$

Then since for all g

$$\int g^*\chi\big(\big|\dot{\nabla}^{\hat{E}_\flat}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{s}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\dot{\partial}^{\boldsymbol{t}}\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2+\big|\boldsymbol{x}_\alpha^*(\theta_\alpha\psi)\big|^2\big)=\|\psi\|_{\mathrm{HS}\,1},$$

equation (14) follows.

Using the same arguments with $\dot{\partial}^\alpha$ in place of $\dot{\nabla}^\alpha$, one gets (15).

As for the last inequality, since $\boldsymbol{t}^*E|_{M_\alpha\times\{z\}}$ and the connection $(\boldsymbol{x}_\alpha^{-1})^*\nabla^{\boldsymbol{s}^*E'}$ are trivial along $\exp tZ_0$, one can write

$$\nabla^\alpha_{Z_0}(\hat{A}u\otimes\boldsymbol{s}^*e)=\frac{d}{dt}\Big|_{t=0}Au\Big|_{M_\alpha\times\{\exp tZ\}}\otimes\boldsymbol{s}^*e+u\otimes\nabla^{\boldsymbol{s}^*E'}_{Z_0}\boldsymbol{s}^*e$$
$$=A\bigg(\frac{d}{dt}\Big|_{t=0}u\Big|_{M_\alpha\times\{\exp tZ\}}\bigg)\otimes\boldsymbol{s}^*e+u\otimes\nabla^{\boldsymbol{s}^*E'}_{Z_0}\boldsymbol{s}^*e=\hat{A}\big(\nabla^\alpha_{Z_0}(u\otimes\boldsymbol{s}^*e)\big).$$

It follows that

$$\dot{\partial}^Z\hat{A}\psi^\alpha=\hat{A}(\dot{\partial}^Z\psi^\alpha),$$

from which (16) follows. □

Clearly, the arguments leading to Theorem 2.17 can be repeated and we obtain:

Corollary 2.18. *For any smooth bounded operator A and $m=0,1,\ldots$, there exists $C'_{m,l}>0$ such that for any $\psi\in\Psi^{-\infty}_\infty(M\times_B M)^G$ one has*

$$\|\hat{A}\psi\|_{\mathrm{HS}\,m}\le\bigg(\sum_{0\le l\le m}C_{m,l}\|A\|_{\mathrm{op}\,l}\bigg)\|\psi\|_{\mathrm{HS}\,m}.$$

Notation 2.19. In view of Corollary 2.18, we shall denote

$$\|A\|_{\mathrm{op}'m}:=\bigg(\sum_{0\le l\le m}C_{m,l}\|A\|_{\mathrm{op}\,l}\bigg).$$

We may assume without loss of generality that $C_{m,l}\ge 2$. Then one still has

$$\|A_1A_2\|_{\mathrm{op}'m}\le\|A_1\|_{\mathrm{op}'m}\|A_2\|_{\mathrm{op}'m}. \tag{17}$$

3. Large time behavior of the heat operator

In this section we will prove that under the condition of the positivity of the Novikov–Shubin invariant, the heat operator also convergences to the projection operator under the norm $\|\cdot\|_{\mathrm{HS}\,m}$.

3A. ***The Novikov–Shubin invariant.*** Let $M \to B$ be a fiber bundle with a G action, and $TM = H \oplus V$ be the G-invariant splitting, as defined in Section 2A. Recall that we assumed the metric on $H \cong \pi^*TB$ is given by pulling back some Riemannian metric on B. In other words, V is a Riemannian foliation.

Let $E \to M$ be a flat, contravariant G-vector bundle, and ∇ be an invariant flat connection on E. Denote $E^\bullet := \wedge^\bullet V' \otimes E$.

Since the vertical distribution V is integrable, the de Rham differential $d_V^{\nabla^E}$ along V is well defined. Write $\eth_0 := d_V^{\nabla^E} + \big(d_V^{\nabla^E}\big)^*$, $\Delta := \eth_0^2$, and denote by $e^{-t\Delta}$ the heat operator and Π_0 the orthogonal projection onto $\mathrm{Ker}(\Delta)$.

The following result is classical: See, for example, [Bismut 1986, Proposition 2.8] and [Heitsch 1995, Proposition 3.5].

Lemma 3.1. *The heat operator $e^{-t\Delta}$ is given by a smooth kernel. Moreover, for any first order differential operator A, one has the Duhamel type formula*

$$[A, e^{-t\Delta}] = -\int_0^t e^{-(t-t')\Delta}[A, \Delta]e^{-t'\Delta}\, dt'. \tag{18}$$

From Lemma 3.1, it follows that:

Corollary 3.2 [Heitsch 1995, Corollary 3.11]. *For any i, j, k, there exist $C, M > 0$ such that*

$$\big|(\dot{\nabla}^{E_\flat})^i(\dot{\partial}^{s})^j(\dot{\partial}^{t})^k e^{-t\eth_0^2}\big|(x, y, z) \le Ce^{-M\boldsymbol{d}(y,z)^2}.$$

Hence $e^{-t\eth_0^2} \in \Psi_\infty^{-\infty}(M \times_B M, E^\bullet)^G$.

As for Π_0, one has

Lemma 3.3. *The kernel of Π_0 lies in $\bar{\Psi}_0^{-\infty}(M \times_B M, E^\bullet)^G$.*

Proof. By [Gong and Rothenberg 1996, Theorem 2.2] Π_0 is also represented by a smooth kernel $\Pi_0(x, y, z)$. Moreover by the same theorem and the fact that $\Pi_0 = \Pi_0^2$, one has

$$\sup_{x\in B}\left\{\int_{Z_x} \chi(x, z)\int_{Z_x} |\Pi_0(x, y, z)|^2 \mu_x(y)\mu_x(z)\right\} = \|\Pi_0\|_\tau < \infty,$$

where $\|\cdot\|_\tau$ is the τ-trace norm defined in [Gong and Rothenberg 1996] (see also [Azzali et al. 2015]).

Hence we are left to consider $\chi_n(x, y, z)\Pi_0(x, y, z)$, where $\chi_n \in C^\infty(M \times_B M)^G$ is a sequence of smooth functions such that

(1) $0 \leq \chi_n \leq 1$;

(2) χ_n is increasing and converges pointwise to 1;

(3) $\chi_n(x, y, z) = 0$ whenever $\boldsymbol{d}(y, z) > nr$ for some $r > 0$.

To construct χ_n, let $r > 0$ to be the infimum of the injective radius of the fibers Z_x, and ϕ_1 be a nonnegative smooth function such that $\phi_1(t) = 1$ if $t < \frac{1}{2}r$, $\phi_1(t) = 0$ if $t > r$. Then $\chi_1 := \phi_1 \circ \boldsymbol{d}(y, z)$ is G-invariant. Define

$$\tilde{\chi}_n := \chi_1 \star \cdots \star \chi_1 \text{ (convolution by } n \text{ times).}$$

Note that $\tilde{\chi}_n(x, y, z) > 0$ whenever $\boldsymbol{d}(y, z) < \frac{1}{2}nr$. Moreover, $\tilde{\chi}_n$ is G-invariant and $\tilde{\chi}_n(x, y, z) = 0$ whenever $\boldsymbol{d}(y, z) > nr$. Since $\tilde{\chi}_{n+1}$ is bounded away from 0 on the support of $\tilde{\chi}_n$, clearly one can find smooth functions ϕ_n such that $\chi_n := \phi_n \circ \tilde{\chi}_n$ satisfies conditions (1)–(3). □

Because of Corollary 3.2 and Lemma 3.3, it makes sense to define:

Definition 3.4. We say that Δ has positive Novikov–Shubin invariant if there exist $\gamma > 0$ and $C_0 > 0$ such that for sufficiently large t,

$$\sup_{x \in B} \left\{ \int_{Z_x} \chi(x, z) \int_{Z_x} \left| (e^{-t\Delta} - \Pi_0)(x, y, z) \right|^2 \mu_x(y) \mu_x(z) \right\} \leq C_0 t^{-\gamma}.$$

Remark 3.5. The positivity of the Novikov–Shubin invariant is independent of the metrics defining the operator Δ.

Remark 3.6. Since $e^{-(t/2)\Delta} - \Pi_0$ is nonnegative, self adjoint and $(e^{-(t/2)\Delta} - \Pi_0)^2 = e^{-t\Delta} - \Pi_0$, one has

$$\sup_{x \in B} \left\{ \int_{Z_x} \chi(x, z) \int_{Z_x} \left| (e^{-\frac{t}{2}\Delta} - \Pi_0)(x, y, z) \right|^2 \mu_x(y) \mu_x(z) \right\} = \| e^{-t\Delta} - \Pi_0 \|_\tau.$$

Hence our definition of having positive Novikov–Shubin invariant is equivalent to that of [Azzali et al. 2015]. Our argument here is similar to the proof of [Bismut et al. 2017, Theorem 7.7].

In this paper, we shall always assume Δ has positive Novikov–Shubin invariant. From this assumption, it follows by integration over B that there exist constants $\gamma > 0$ and $C > 0$ such that for t large enough

$$\| e^{-t\Delta} - \Pi_0 \|_{\mathrm{HS}\,0} < C t^{-\gamma}. \tag{19}$$

3B. *Example: The Bismut superconnection.*

Definition 3.7. A standard flat Bismut superconnection is an operator of the form

$$d^{\nabla^E} := d_V^{\nabla^E} + \nabla^{E_\flat^\bullet} + \iota_\Theta,$$

where Θ is the V-valued horizontal 2-form defined by

$$\Theta(X_1^H, X_2^H) := -P^V[X_1^H, X_2^H] \quad \text{for all } X_1, X_2 \in \Gamma^\infty(TB),$$

and ι_Θ is the contraction with Θ. Note that P^V is not canonical and it depends on the splitting $TM = V \oplus H$.

Observe that the adjoint of the Bismut superconnection,

$$\big(d^{\nabla^E}\big)' = \big(d_V^{\nabla^E}\big)^* + \big(\nabla^{E_\flat^\bullet}\big)' - \Lambda_{\Theta^*},$$

is also flat. It follows that

$$\big(\nabla^{E_\flat^\bullet}\big)'\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^*\big(\nabla^{E_\flat^\bullet}\big)' = 0.$$

Define

$$\Omega := \tfrac{1}{2}\Big(\big(\nabla^{E_\flat^\bullet}\big)' - \nabla^{E_\flat^\bullet}\Big).$$

Observe that Ω is a tensor (see [Álvarez López and Kordyukov 2001] for an explicit formula for Ω). Moreover one has

$$\nabla^{E_\flat^\bullet}\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^*\nabla^{E_\flat^\bullet} = 2\Omega\big(d_V^{\nabla^E}\big)^* + 2\big(d_V^{\nabla^E}\big)^*\Omega.$$

Also, observe that $\big(d_V^{\nabla^E}\big) + \big(d_V^{\nabla^E}\big)^* + \nabla^{E_\flat^\bullet} + \Big(\big(\nabla^{E_\flat^\bullet}\big)'\Big)^*$ is an elliptic operator.

3C. ***The regularity result of Alvarez Lopez and Kordyukov.*** We first recall that an operator A is called C^∞-bounded if in normal coordinates the coefficients and their derivatives are uniformly bounded. As in [Álvarez López and Kordyukov 2001], we make the more general assumption that there exists C^∞-bounded first order differential operator Q, and zero degree operators R_1, R_2, R_3, R_4, all G-invariant, such that $d_V^{\nabla^E} + \big(d_V^{\nabla^E}\big)^* + Q$ is elliptic, and

$$\begin{aligned} Q\, d_V^{\nabla^E} + d_V^{\nabla^E} Q &= R_1\, d_V^{\nabla^E} + d_V^{\nabla^E} R_2, \\ Q\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^* Q &= R_3\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^* R_4. \end{aligned} \tag{20}$$

Clearly, in our example, $\nabla^{E_\flat^\bullet} + \Big(\big(\nabla^{E_\flat^\bullet}\big)'\Big)^*$ satisfies Equation (20).

Write $\eth_0 := d_V^{\nabla^E} + (d_V^{\nabla^E})^*$, $\Delta := \eth_0^2$, and denote by Π_{d_V} and $\Pi_{d_V^*}$ respectively the orthogonal projections onto the range of $d_V^{\nabla^E}$ and $\big(d_V^{\nabla^E}\big)^*$, which we shall denote by $\mathrm{Rg}(d_V)$ and $\mathrm{Rg}(d_V^*)$.

In this section, we shall consider the operators

$$B_1 := R_1\Pi_{d_V} + R_3\Pi_{d_V^*}, \quad B_2 := \Pi_{d_V^*} R_2 + \Pi_{d_V} R_4, \quad B := B_2\Pi_0 + B_1(\mathrm{id} - \Pi_0).$$

We recall some elementary formulas regarding these operators from [Álvarez López and Kordyukov 2001]:

Lemma 3.8 [Álvarez López and Kordyukov 2001, Lemma 2.2]. *One has*

$$\begin{aligned} Q\, d_V^{\nabla^E} + d_V^{\nabla^E} Q &= B_1 d_V^{\nabla^E} + d_V^{\nabla^E} B_2, \\ Q\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^* Q &= B_1\big(d_V^{\nabla^E}\big)^* + \big(d_V^{\nabla^E}\big)^* B_2, \\ [Q, \Delta] &= B_1 \Delta - \Delta B_2 - \eth_0(B_1 - B_2)\eth_0. \end{aligned}$$

One can furthermore estimate the derivatives of Π_0. First, recall that

Lemma 3.9. *One has* (*see* [Álvarez López and Kordyukov 2001, Corollary 2.8])

$$[Q + B, \Pi_0] = 0.$$

Proof. Here we give a different proof. From definition we have

$$B = (\Pi_{d_V^*} R_2 + \Pi_d R_4)\Pi_0 + R_1 \Pi_{d_V} + R_3 \Pi_{d_V^*},$$

where we used $\Pi_{d_V} \Pi_0 = \Pi_{d_V^*} \Pi_0 = 0$. Hence

$$B\Pi_0 - \Pi_0 B = (\Pi_{d_V^*} R_2 + \Pi_{d_V} R_4)\Pi_0 - \Pi_0 R_1 \Pi_{d_V} - \Pi_0 R_3 \Pi_{d_V^*}.$$

For any s one has

$$\Pi_{d_V} s = \lim_{n\to\infty} d\tilde{s}_n,$$

for some sequence $\tilde{s}_n$ (in some suitable function spaces). It follows that

$$\Pi_0 R_1 \Pi_{d_V} s = \lim_{n\to\infty} \Pi_0 R_1 d\tilde{s}_1 = \lim_{n\to\infty} \Pi_0\big(Q d_V^{\nabla^E} + d_V^{\nabla^E} Q - d_V^{\nabla^E} R_2\big)\tilde{s}_1 = \Pi_0 Q \Pi_{d_V} s.$$

Similarly, one has $\Pi_0 R_3 \Pi_{d_V^*} = \Pi_0 Q \Pi_{d_V^*}$ and by considering the adjoint,

$$\Pi_{d_V^*} R_2 \Pi_0 = \Pi_{d_V^*} Q \Pi_0 \quad \text{and} \quad \Pi_{d_V^*} R_4 \Pi_0 = \Pi_{d_V^*} Q \Pi_0.$$

It follows that

$$[Q + B, \Pi_0] = (\mathrm{id} - \Pi_{d_V} - \Pi_{d_V^*}) Q \Pi_0 - \Pi_0 Q (\mathrm{id} - \Pi_{d_V} - \Pi_{d_V^*}) = 0. \qquad \square$$

In other words, regarding $[Q, \Pi_0]$ and $[B, \Pi_0]$ as kernels, one has

$$\|[Q, \Pi_0]\|_{\mathrm{HS}m} = \|[B, \Pi_0]\|_{\mathrm{HS}m},$$

provided the right-hand side is finite. Hence, using elliptic regularity and the same arguments as Lemma 3.3, one can prove inductively that

$$\Pi_0(x, y, z) \in \bar{\Psi}_m^{-\infty}(M \times_B M, E^\bullet) \quad \text{for all } m.$$

Next, we recall the main result of [Álvarez López and Kordyukov 2001]

Lemma 3.10. *For any* $m = 0, 1, \dots,$

(1) *The heat operator* $e^{-t\Delta}$, *and the operators* $\eth_0 e^{-t\Delta}$, $\Delta e^{-t\Delta}$ *map* $\mathcal{W}^m(E)$ *to itself as bounded operators. Moreover, there exist constants* $C_m^0, C_m^1, C_m^2 > 0$ *such that*

$$\|e^{-t\Delta}\|_{\mathrm{op}\,m} \le C_m^0, \quad \|\eth_0 e^{-t\Delta}\|_{\mathrm{op}\,m} \le t^{-\frac{1}{2}} C_m^1, \quad \|\Delta e^{-t\Delta}\|_{\mathrm{op}\,m} \le t^{-1} C_m^2,$$

for all $t > 0$.

(2) *As* $t \to \infty$, $e^{-t\Delta}$ *strongly converges as an operator on* $\mathcal{W}^m(E)$. *Moreover*, $(t, s) \mapsto e^{-t\Delta}s$ *is a continuous map from* $[0, \infty] \times \mathcal{W}^m(E)$ *to* $\mathcal{W}^m(E)$.

(3) *One has the Hodge decomposition*

$$\mathcal{W}^m(E) = \mathrm{Ker}(\Delta) + \overline{\mathrm{Rg}(\Delta)} = \mathrm{Ker}(\eth_0) + \overline{\mathrm{Rg}(\eth_0)},$$

where the kernel, image and closure are in $\mathcal{W}^m(E)$.

Note that our case is slightly different from that of [Álvarez López and Kordyukov 2001], where M is assumed to be compact (but with possibly noncompact fibers). However, the same arguments clearly apply because our M is of bounded geometry.

We recall more results in [Álvarez López and Kordyukov 2001, Section 2].

Lemma 3.11 [Álvarez López and Kordyukov 2001, Lemma 2.4]. *For any* $m \ge 0$, *there exists a constant* $C_m^3 > 0$ *such that*

$$\|[Q, e^{-t\Delta}]\|_{\mathrm{op}\,m} \le C_m^3.$$

Proof. Using the third equation of Lemma 3.8, equation (18) becomes

$$[Q, e^{-t\Delta}] = \int_0^t e^{-(t-t')\Delta}\eth_0(B_1 - B_2)\eth_0 e^{-t'\Delta}\,dt' - \int_0^t e^{-(t-t')\Delta}(B_1\Delta - \Delta B_2)e^{-t'\Delta}\,dt'.$$

Using Lemma 3.10, we estimate the first integral

$$\begin{aligned}\left\| \int_0^t e^{-(t-t')\Delta}\eth_0(B_1 - B_2)\eth_0 e^{-t'\Delta}\,dt' \right\|_{\mathrm{op}\,m} &\le \|B_1 - B_2\|_{\mathrm{op}\,m}(C_m^1)^2 \int_0^t \frac{dt'}{\sqrt{(t-t')t'}} \\ &= \|B_1 - B_2\|_{\mathrm{op}\,m}(C_m^1)^2\pi.\end{aligned}$$

As for the second integral, we split the domain of integration into $\left[0, \frac{1}{2}t\right]$ and $\left[\frac{1}{2}t, t\right]$, and then integrate by part to get

$$\begin{aligned}&\int_0^t e^{-(t-t')\Delta}(B_1\Delta - \Delta B_2)e^{-t'\Delta}\,dt' \\ &\quad = \int_0^{t/2} e^{-(t-t')\Delta}\Delta(-B_1 - B_2)e^{-t'\Delta}\,dt' - \int_{t/2}^t e^{-(t-t')\Delta}(B_1 - B_2)\Delta e^{-t'\Delta}\,dt' \\ &\qquad + e^{-(t-t')\Delta}B_1 e^{-t'\Delta}\Big|_{t'=0}^{t/2} - e^{-(t-t')\Delta}B_2 e^{-t'\Delta}\Big|_{t'=t/2}^{t}.\end{aligned}$$

Again using Lemma 3.10, its $\|\cdot\|_{\mathrm{op}\,m}$-norm is bounded by

$$C^0_m C^1_m(\|B_1\|_{\mathrm{op}\,m}+\|B_2\|_{\mathrm{op}\,m})\biggl(\int_0^{t/2}\frac{dt'}{t-t'}+\int_{t/2}^{t}\frac{dt'}{t'}\biggr)+C^0_m(C^0_m+1)(\|B_1\|_{\mathrm{op}\,m}+\|B_2\|_{\mathrm{op}\,m}),$$

which is uniformly bounded because $\int_0^{t/2} 1/(t-t')\,dt' = \int_{t/2}^{t} 1/t'\,dt' = \log 2$. $\square$

Lemma 3.9 suggests that $[Q+B, e^{-t\Delta}]$ converges to zero as $t\to\infty$. Indeed, we shall prove a stronger result, namely, $[Q+B, e^{-t\Delta}]$ decays polynomially in the $\|\cdot\|_{\mathrm{HS}\,m}$-norm for all m.

Lemma 3.12. *Suppose there exist $C_m, \gamma>0$ such that $\|e^{-t\Delta}-\Pi_0\|_{\mathrm{HS}\,m}\le C_m t^{-\gamma}$, then there exist $C'_m, \gamma_m>0$ such that*

$$\|[Q+B, e^{-t\Delta}]\|_{\mathrm{HS}\,m}=\|[Q+B, e^{-t\Delta}-\Pi_0]\|_{\mathrm{HS}\,m}\le C'_m t^{-\gamma_m}.$$

Proof. We follow the proof of [Álvarez López and Kordyukov 2001, Lemma 2.6]. By Lemma 3.8, we get

$$[Q+B,\Delta]=(\Delta(B_1+B_2)+\eth_0(B_1-B_2)\eth_0)(\mathrm{id}-\Pi_0).$$

It follows that $\Pi_0[Q+B, e^{-(t/2)\Delta}]=[Q+B, e^{-(t/2)\Delta}]\Pi_0=0$. Write

$$\begin{aligned}[Q+B, e^{-t\Delta}]&=[Q+B, e^{-\frac12 t\Delta}]e^{-\frac12 t\Delta}+e^{-\frac12 t\Delta}[Q+B, e^{-\frac12 t\Delta}]\\&=[Q+B, e^{-\frac12 t\Delta}](e^{-\frac12 t\Delta}-\Pi_0)+(e^{-\frac12 t\Delta}-\Pi_0)[Q+B, e^{-\frac12 t\Delta}].\end{aligned}$$

Taking $\|\cdot\|_{\mathrm{HS}\,m}$ and using Corollary 2.18 and Lemma 3.11, the claim follows. $\square$

Theorem 3.13. *Suppose $\|e^{-t\Delta}-\Pi_0\|_{\mathrm{HS}\,0}\le C_0 t^{-\gamma}$ for some $\gamma>0$, $C_0>0$. Then for any m, there exists $C''_m>0$ such that*

$$\|e^{-t\Delta}-\Pi_0\|_{\mathrm{HS}\,m}\le C''_m t^{-\gamma}\quad \textit{for all } t>1.$$

Proof. We prove the theorem by induction. The case $m=0$ is given. Suppose that for some m, $\|e^{-t\Delta}-\Pi_0\|_{\mathrm{HS}\,m}\le C_m t^{-\gamma}$. Consider $\|e^{-t\Delta}-\Pi_0\|_{\mathrm{HS}\,m+1}$.

Since Q is a first order differential operator, for any kernel $\psi\in\Psi^{-\infty}_{\infty}(M\times_B M, E^\bullet)^G$, $[Q,\psi]$ is also a kernel lying in $\Psi^{-\infty}_{\infty}(M\times_B M, E^\bullet)^G$, that is in particular, given by a composition of the covariant derivatives $\dot\nabla^{\hat E_\flat}$, $\dot\partial^s$, $\dot\partial^t$ and some tensors acting on ψ. Since $\|\psi\|_{\mathrm{HS}\,m}$ is by definition the $\|\cdot\|_{\mathrm{HS}\,0}$ norm of the m-th derivatives of ψ, elliptic regularity (Lemma 2.12) implies

$$\|\psi\|_{\mathrm{HS}\,m+1}\le \tilde C_m(\|\psi\|_{\mathrm{HS}\,m}+\|\eth_0\psi\|_{\mathrm{HS}\,m}+\|\psi\eth_0\|_{\mathrm{HS}\,m}+\|[Q,\psi]\|_{\mathrm{HS}\,m}),$$

for some constant $\tilde{C}_m > 0$. Put $\psi = e^{-t\Delta} - \Pi_0$. The theorem then follows from the estimates

$$
\begin{aligned}
\|\eth_0(e^{-t\Delta}-\Pi_0)\|_{\mathrm{HS}\,m} &= \|(e^{-t\Delta}-\Pi_0)\eth_0\|_{\mathrm{HS}\,m} \\
&\le \Big(\sum_{0\le l\le m} C'_{m,l}\|\eth_0(e^{-(t/2)\Delta}-\Pi_0)\|_{\mathrm{op}\,l}\Big)\|e^{-(t/2)\Delta}-\Pi_0\|_{\mathrm{HS}\,m} \\
&\le \Big(\sum_{0\le l\le m} C'_{m,l}C_l^1\big(\tfrac{1}{2}t\big)^{-1/2}\Big)C_m\big(\tfrac{1}{2}t\big)^{-\gamma}, \\
\|[Q, e^{-t\Delta}-\Pi_0]\|_{\mathrm{HS}\,m} &\le \|[Q+B, e^{-t\Delta}-\Pi_0]\|_{\mathrm{HS}\,m} + \|[B, e^{-t\Delta}-\Pi_0]\|_{\mathrm{HS}\,m} \\
&\le C'_m t^{-\gamma} + 2\Big(\sum_{0\le l\le m} C'_{m,l}\|B\|_{\mathrm{op}\,l}\Big)C_m t^{-\gamma}.
\end{aligned}
$$

Note that we used Lemma 3.12 for the last inequality. □

4. Sobolev convergence

In this section we will use the method of [Azzali et al. 2015] to prove that under the condition of positivity of the Novikov–Shubin invariant the L^2-analytic torsion form is a smooth form.

Let ∇^E be a flat connection on E. Define the number operators on $\wedge^\bullet H'\otimes\wedge^\bullet V'\otimes E$ by

$$N_\Omega|_{\wedge^q H'\otimes\wedge^{q'}V'\otimes E} := q \quad\text{and}\quad N|_{\wedge^q H'\otimes\wedge^{q'}V'\otimes E} := q'.$$

In this section, we consider the rescaled Bismut superconnection [Berline et al. 1992, Chapter 9.1]

$$
\begin{aligned}
\eth(t) &:= \tfrac{1}{2}t^{1/2}t^{-N_\Omega/2}(d+d^*)t^{N_\Omega/2} \\
&= \tfrac{1}{2}\big(t^{1/2}(d_V+d_V^*) + \big(\nabla^{E_\flat}+(\nabla^{E_\flat})'\big) + t^{-1/2}(-\Lambda_{\Theta^*}+\iota_\Theta)\big).
\end{aligned}
$$

Denote

$$
\begin{aligned}
D_0 &:= -\tfrac{1}{2}(d_V-d_V^*), \quad \Omega_t := -\tfrac{1}{2}\big(\nabla^{E_\flat}-(\nabla^{E_\flat})'\big) - \tfrac{1}{2}t^{-1/2}(-\Lambda_{\Theta^*}-\iota_\Theta), \\
D(t) &:= t^{1/2}D_0+\Omega_t.
\end{aligned}
$$

The curvature of $\eth(t)$ can be expanded in the form:

$$\eth(t)^2 = -D(t)^2 = t\Delta + t^{1/2}\Omega_t D_0 + t^{1/2}D_0\Omega_t + \Omega_t^2.$$

Hence as a consequence of Duhamel's expansion (see [Berline et al. 1992]), we have

$$e^{-\eth(t)^2} = e^{D(t)^2} = e^{-t\Delta} + \sum_{n=1}^{\dim B} \int_{(r_0,\dots,r_k)\in\Sigma^n} e^{-r_0 t\Delta}(t^{1/2}\Omega_t D_0 + t^{1/2} D_0 \Omega_t + \Omega_t^2) e^{-r_1 t\Delta} \cdots (t^{1/2}\Omega_t D_0 + t^{1/2} D_0 \Omega_t + \Omega_t^2) e^{-r_n t\Delta}\, d\Sigma^n,$$

where $\Sigma^n := \{(r_0, r_1, \dots, r_n) \in [0,1]^{n+1} : r_0 + \cdots + r_n = 1\}$.

4A. ***The large time estimate of the rescaled heat operator.*** In this section, we follow [Azzali et al. 2015, Section 4] to estimate the Hilbert–Schmidt norms of $e^{-\eth(t)^2}$ (see Theorem 4.4 below).

Let $\gamma' := 1 - (1 + 2\gamma/(\dim B + 2 + 2\gamma))^{-1}$, $\bar{r}(t) := t^{-\gamma'}$. Fix $\bar{t}$ such that $\bar{r}(\bar{t}) < (\dim B + 1)^{-1}$. One has the following counterparts of [Azzali et al. 2015, Lemma 4.2]:

Lemma 4.1. *For $c = 0, 1, 2$, there exists a constant C_m such that*

$$\big\|(\sqrt{t}\eth_0)^{c/2} e^{rt(D_0)^2}\big\|_{\mathrm{op}'m} \le C_m r^{-c/2} \text{ for any } t > \bar{t},\ 0 < r < 1 \text{ (by Lemma 3.10)};$$

and for any $t > \bar{t}$, $\bar{r}(t) < r < 1$,

$$\begin{aligned} \big\|e^{rt(D_0)^2}\big\|_{\mathrm{HS}m} &\le C_m (rt)^{-\gamma} && \text{(by Theorem 3.13)},\\ \big\|(\sqrt{t}\eth_0)^{c/2} e^{rt(D_0)^2}\big\|_{\mathrm{HS}m} &\le C_m r^{-c/2}(rt)^{-\gamma} \text{ if } c = 1, 2 && \text{(by Corollary 2.18)}. \end{aligned}$$

We furthermore observe that the arguments leading to the main result [Azzali et al. 2015, Theorem 4.1] still hold if one replaces the operator and $\|\cdot\|_\tau$ norm respectively by $\|\cdot\|_{\mathrm{op}'m}$ and $\|\cdot\|_{\mathrm{HS}m}$ for any m.

The arguments in [Azzali et al. 2015, Section 4] are elementary, so we shall only recall some key steps.

First, one splits the domain of integration $\Sigma^n = \bigcup_{I \neq \{0,\dots,n\}} \Sigma^n_{\bar{r}(t),I}$, where

$$\Sigma^n_{\bar{r}(t),I} := \{(r_0,\dots,r_n) : r_i \le \bar{r}(t), \text{ for all } i \in I, r_j \ge \bar{r}(t), \text{ for all } j \notin I\}.$$

Define

$$\begin{aligned} &K(t, n, I, c_0, \dots, c_n; a_1, \dots a_n) \\ &\quad := \int_{\Sigma^n_{\bar{r}(t),I}} (t^{1/2} D_0)^{c_0} e^{-r_0 t\Delta} \prod_{i=1}^n \big(\Theta_t^{a_i} (t^{1/2} D_0)^{c_i} e^{-r_i t\Delta}\big)\, d\Sigma^n, \end{aligned} \tag{21}$$

for $c_i = 0, 1, 2$, $a_j = 1, 2$. Then one has

$$e^{-\eth(t)^2} = e^{D(t)^2} = \sum K(t, n, I, c_0, \dots, c_n; a_1, \dots, a_n),$$

by grouping terms involving D_0 together.

We shall consider the kernels $K(t, n, I, c_0, \dots, c_n; a_1, \cdots, a_n)(x, y, z)$ of the terms in the summation above. Consider the special case when $c_i = 0, 1$. One has the analogue of [Azzali et al. 2015, Proposition 4.6]:

Lemma 4.2. *There exists $\varepsilon > 0$ such that as $t \to \infty$,*

$$K(t, n, I, c_0, \dots, c_n, a_1, \dots, a_n)(x, y, z) = \begin{cases} \left(\frac{\Pi_0 \Omega^{a_1} \Pi_0 \cdots \Pi_0}{n!}\right)(x, y, z) + O(t^{-\varepsilon}) & \text{if } I = \varnothing, \text{ all } c_i = 0 \\ O(t^{-\varepsilon}) & \text{otherwise} \end{cases}$$

in the $\|\cdot\|_{\mathrm{HS}m}$-norm.

Proof. We first consider the case $I = \varnothing$. Suppose furthermore $c_q = 1$ for some q. By Corollary 2.18, The $\|\cdot\|_{\mathrm{HS}m}$-norm of the integrand on the right-hand side of (21) is bounded by

$$\begin{aligned} \left\|(t^{1/2} D_0)^{c_0} e^{-r_0 t\Delta}\right\|_{\mathrm{op}'m} \cdots \left\|\Omega_t^{a_q}\right\|_{\mathrm{op}'m} \left\|(t^{1/2} D_0) e^{-r_q t\Delta}\right\|_{\mathrm{HS}m} \cdots \left\|(t^{1/2} D_0)^{c_n} e^{-r_n t\Delta}\right\|_{\mathrm{op}'m} \\ \leq C'_m r_0^{-c_0/2} \cdots r_q^{-c_q/2} (r_q t)^{-\gamma} \cdots r_n^{-c_n/2} \\ \leq C'_m \bar{r}(t)^{-n/2-\gamma} t^{-\gamma}. \end{aligned}$$

Integrating, we have the estimate

$$\|K(t, n, c_0, \dots, c_n; a_1, \dots, a_n)(x, y, z)\|_{\mathrm{HS}m} \leq C'_m t^{-\gamma + \gamma'(n/2+\gamma)} \int d\Sigma^n,$$

which is $O(t^{-\varepsilon})$ with $\varepsilon = \gamma(1 - (\dim B + 2\gamma)/(\dim B + 2 + 2\gamma))$.

Next, suppose $I = \varnothing$ and $c_i = 0$ for all i. Write $e^{-r_0 t\Delta - \Pi_0} + \Pi_0$ and split the integrand

$$(e^{-r_0 t\Delta} \Omega_t^{a_1} e^{-r_1 t\Delta} \cdots e^{-r_n t\Delta})(x, y, z)$$

into 2^{n+1} terms. If any term contains a $e^{-r_i t\Delta} - \Pi_0$ factor, similar arguments as above shows that it is $O(t^{-\gamma})$. Hence the only term that does not converge to 0 is

$$(\Pi_0 \Omega^{a_1} \Pi_0 \cdots \Pi_0)(x, y, z).$$

Since the volume of $\Sigma^n_{\bar{r}(t), I}$ converges to $\frac{1}{n!}$ as $t \to \infty$, the claim follows.

We are left to consider the case when I is nonempty. Write $I = \{i_1, \dots, i_s\}$, $\{0, \dots, n\} \setminus I =: \{k_1, \dots, k_{s'}\} \neq \varnothing$. For t sufficiently large $I \neq \{0, \dots, n\}$. Suppose $c_q = 1$ for some $q \notin I$. Then we take $\|\cdot\|_{\mathrm{HS}m}$-norm for $(t^{1/2} D_0) e^{-r_q t\Delta}$ term, and estimate

$$\begin{aligned}&\|K(t,n,c_0,\dots,c_n;a_1,\dots,a_n)(x,y,z)\|_{\mathrm{HS}m}\\ &\qquad\le\int_0^{\bar r(t)}\cdots\int_0^{\bar r(t)}\Big(\int_{\{(r_{k_1},\dots,r_{k'_s}):(r_0,\dots,r_n)\in\Sigma^n_{\bar r(t),I}\}}C'_m r_0^{-c_0/2}\cdots r_q^{-c_q/2}(r_qt)^{-\gamma}\cdots r_n^{-c_n/2}\\ &\hspace{22em} d(r_{k_1}\cdots r_{k_{s'}})\Big)dr_{i_1}\cdots dr_{i_s}.\end{aligned}$$

As in the $I=\varnothing$ case, the integral over $\{(r_{k_1}),\dots,r_{k_{s'}}:(r_0,\dots,r_n)\in\Sigma^n_{\bar r(t),I}\}$ is $O(t^{-\varepsilon})$, while $\int_0^{\bar r(t)} r_i^{c_i/2}\,dr_i=O(t^{-\gamma'(1-c_i/2)})$. Again the claim is verified.

Finally if $c_i=0$ for all $i\in I$, then

$$\begin{aligned}&\|K(t,n,c_0,\dots,c_n;a_1,\dots,a_n)(x,y,z)\|_{\mathrm{HS}m}\\ &\le\int_0^{\bar r(t)}\cdots\int_0^{\bar r(t)}\Big(\int_{\{(r_{k_1},\dots,r_{k'_s}):(r_0,\dots,r_n)\in\Sigma^n_{\bar r(t),I}\}}C''_m r_0^{-c_{i_1}/2}\cdots r_n^{-c_{i_s}/2}d(r_{k_1}\cdots r_{k_{s'}})\Big)dr_{i_1}\cdots dr_{i_s}\\ &=O(t^{-\gamma'(1-c_i/2)}).\end{aligned}$$

□

One then turns to the case for some i, $c_i=2$. If I and J are disjoint subsets of $\{0,\dots,n\}$ with $I=\{i_1,\dots,i_r\}$, and $\{0,\dots,n\}\setminus(I\cup J)=:\{k_0,\dots,k_q\}\neq\varnothing$, denote

$$\Sigma^n_{\bar r(t),I,J}:=\{(r_0,\dots,r_n)\in\Sigma^n_{\bar r(t),I}:r_j=\bar r(t),\ \text{whenever } j\in J\},$$

and define

$$\begin{aligned}K(t,n,I,J,c_0,\dots,c_n;a_1,\dots,a_n):=&\int_0^{\bar r(t)}\cdots\int_0^{\bar r(t)}\int_{\{(r_{k_0},\dots r_{k_q}):(r_0,\dots,r_n)\in\Sigma^n_{\bar r(t),I}\}}(t^{1/2}D_0)^{c_0}e^{-r_0t\Delta}\\ &\prod_{i=1}^n\big(\Theta_t^{a_i}(t^{1/2}D_0)^{c_i}e^{-r_it\Delta}\big)\big|_{\Sigma^n_{\bar r(t),I,J}}d^q(r_{k_0},\dots,r_{k_q})dr_1\cdots dr_r.\end{aligned}$$

Using integration by parts, one gets [Azzali et al. 2015, Equation (4.17)],

$$\begin{aligned}&K(t,n,I\cup\{i_p\},J;\dots,2,\dots,c_{k_0},\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\\ &=\begin{cases}K(t,n,I,J\cup\{i_p\};\dots,0,\dots,c_{k_0},\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\\ -K(t,n-1,I,J;\dots,\dots,c_{k_0},\dots;\dots,a_{i_0}+a_{i_{p+1}},\dots) & q>0,\\ +K(t,n,I\cup\{i_p\},J\cup\{k_0\};\dots,0,\dots,c_{k_0},\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\\ +K(t,n,I\cup\{i_p\},J;\dots,0,\dots,c_{k_0}+2,\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\\ K(t,n,I,J\cup\{i_p\};\dots,0,\dots,c_{k_0},\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\\ -K(t,n-1,I,J;\dots,\dots,c_{k_0},\dots;\dots,a_{i_0}+a_{i_{p+1}},\dots) & q=0.\\ +K(t,n,I\cup\{i_p\},J;\dots,0,\dots,c_{k_0}+2,\dots;\dots,a_{i_p},a_{i_{p+1}},\dots)\end{cases}\end{aligned}\tag{22}$$

We remark that the proof of [Azzali et al. 2015, Equation (4.17)] does not involve any norm, and therefore we omit the details here.

Using equation (22) repeatedly, one eliminates all terms with $c_i = 2$.

On the other hand one has the following straightforward generalization of Lemma 4.2 (compare with [Azzali et al. 2015, Proposition 4.7]):

Lemma 4.3. *Suppose $c_i = 0, 1$. As $t \to \infty$,*

$$K(t,n,I,J,c_0,\dots c_n;a_1,\dots,a_n)(x,y,z) \\ = \begin{cases} \bigl(\frac{1}{(n-|J|)!}\Pi_0\Omega^{a_1}\Pi_0\cdots\Pi_0\bigr)(x,y,z) + O(t^{-\gamma'}) & \text{if } I=\varnothing, c_0,\dots,c_n = 0 \\ O(t^{-\gamma'}) & \text{otherwise,} \end{cases}$$

for some $\gamma' > 0$, in the $\|\cdot\|_{\mathrm{HS}m}$-norm.

Thus the term $K(t, n, I, c_0, \dots c_n; a_1, \dots a_n)$ converges to 0 unless

$$c_i = 0 \text{ whenever } i \in I, \quad c_i = 2 \text{ whenever } i \notin I.$$

Then one follows exactly as [Azzali et al. 2015, Section 4.5] to compute the limit, and concludes with the following analogue of their Theorem 4.1:

Theorem 4.4. *For $k = 0, 1, 2$ and any $m \in \mathbb{N}$,*

$$\lim_{t\to\infty} D(t)^k e^{-\eth(t)^2}(x, y, z) = \Pi_0(\Omega\Pi_0)^k e^{(\Omega\Pi_0)^2}(x, y, z)$$

in the $\|\cdot\|_{\mathrm{HS}m}$-norm, where $\Omega := -\frac{1}{2}(\nabla^{E_\flat} - (\nabla^{E_\flat})^*)$. Moreover, there exits $\varepsilon' > 0$ such that as $t \to \infty$,

$$\bigl\|\bigl(D(t)^k e^{-\eth(t)^2} - \Pi_0(\Omega\Pi_0)^k e^{(\Omega\Pi_0)^2}\bigr)(x, y, z)\bigr\|_{\mathrm{HS}m} = O(t^{-\varepsilon'}).$$

4B. *Application: the L^2-analytic torsion form.* Our main application of this theorem is in establishing the smoothness and transgression formula of the L^2-analytic torsion form. Here, we briefly recall the definitions.

On $\wedge^\bullet T^*M \otimes E \cong \wedge^\bullet H' \otimes \wedge^\bullet V' \otimes E$, define N_Ω, N to be the number operators of $\wedge^\bullet H' \cong \pi^{-1}(\wedge^\bullet T^*B)$ and $\wedge^\bullet V'$ respectively.

Define

$$F^\wedge(t) := (2\pi\sqrt{-1})^{-N_\Omega/2}\,\mathrm{str}_\Psi(2^{-1}N(1 + 2D(t)^2)e^{-\eth(t)^2}).$$

Then under the positivity of the Novikov–Shubin invariant, we have the following well-defined L^2-analytic torsion form:

Definition 4.5 [Azzali et al. 2015].

$$\tau := \int_0^\infty \bigl\{-F^\wedge(t) + \tfrac{1}{2}\,\mathrm{str}_\Psi(N\Pi_0) \\ + \Bigl(\tfrac{1}{4}\dim(Z)\,\mathrm{rk}(E)\,\mathrm{str}_\Psi(\Pi_0) - \tfrac{1}{2}\,\mathrm{str}_\Psi(N\Pi_0)\Bigr)(1-2t)e^{-t}\bigr\}\,\frac{dt}{t}.$$

In [Azzali et al. 2015], it is only shown that the form τ is continuous. Next we will show that indeed the form τ is smooth.

Theorem 4.6. *The form τ is smooth, i.e., $\tau \in \Gamma^\infty(\wedge^\bullet T^*B)$.*

Proof. Using [Berline et al. 1992, Proposition 9.24], the derivatives of the t-integrand are bounded as $t \to 0$. It follows that its integral over $[0, 1]$ is smooth.

We turn to studying the large time behavior. Consider $\mathrm{str}(2^{-1}N(e^{-\eth(t)^2} - \Pi_0))$. Using the semigroup property, we can write $e^{-\eth(t)^2} = 2^{-N_\Omega/2}e^{-\eth(t/2)^2}e^{-\eth(t/2)^2}2^{N_\Omega/2}$. Also, since $\mathrm{str}(N\Pi_0(\Omega\Pi_0)^{2j}) = \mathrm{str}([N\Pi_0(\Omega\Pi_0), \Pi_0(\Omega\Pi_0)^{2j-1}]) = 0$ for any $j \geq 1$ one has

$$\mathrm{str}(N\Pi_0) = \mathrm{str}(N\Pi_0 e^{(\Omega\Pi_0)^2}) = 2^{-N_\Omega/2}\,\mathrm{str}(N\Pi_0 e^{(\Omega\Pi_0)^2}\Pi_0 e^{(\Omega\Pi_0)^2}).$$

Therefore

$$\begin{aligned}\mathrm{str}\big(2^{-1}N\big(e^{-\eth(t)^2} - \Pi_0\big)\big) &= 2^{-N_\Omega/2}\mathrm{str}\big(2^{-1}N\big(e^{-\eth(t/2)^2}e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\Pi_0 e^{(\Omega\Pi_0)^2}\big)\big)\\ &= 2^{-N_\Omega/2}\mathrm{str}\big(2^{-1}Ne^{-\eth(t/2)^2}\big(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big)\big)\\ &\qquad + 2^{-N_\Omega/2}\mathrm{str}\big(2^{-1}N\big(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big)\Pi_0 e^{(\Omega\Pi_0)^2}\big).\end{aligned}$$

Now consider the $L^2(B)$-norm of $\mathrm{str}_\Psi\big(2^{-1}Ne^{-\eth(t/2)^2}\big(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big)\big)$. To shorten notations, denote $G := 2^{-1}Ne^{-\eth(t/2)^2}\big(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big)$. Writing G as a convolution product, then there exists a constant $C_0 > 0$ such that

$$\begin{aligned}&\int_B \left|\int_{Z_x} \chi(x,z)\,\mathrm{str}(G(x,z,z))\mu_x(z)\right|^2 \mu_B(x)\\ &= \int_B \left|\int_{Z_x} \chi\,\mathrm{str}\left(\frac{N}{2}\int_{y\in Z_x} e^{-\eth(t/2)^2}(x,z,y)\big(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big)(x,y,z)\mu_x(y)\right)\mu_x(z)\right|^2 \mu_B(x)\\ &\leq C_0\int_B\left(\int_{Z_x}\chi\int_{y\in Z_x}\big|e^{-\eth(t/2)^2}\big|(x,z,y)\big|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big|(x,y,z)\mu_x(y)\mu_x(z)\right)^2\mu_B(x)\\ &\leq C_0\big\|e^{-\eth(t/2)^2}\big\|^2_{\mathrm{HS}\,0}\big\|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\big\|^2_{\mathrm{HS}\,0},\end{aligned}$$

where we used the Cauchy–Schwarz inequality three times. Since $\|e^{-\eth(t/2)^2}\|_{\mathrm{HS}\,0}$ is bounded for t large (by the triangle inequality), the expression above is $O(t^{-\gamma'})$.

We turn to estimating its derivatives. For any vector field X on B,

$$\begin{aligned}\nabla^{TB}_X \mathrm{str}_\Psi(G) = \int&(L_{X^H}\chi(x,z))\,\mathrm{str}(G(x,z,z))\mu_x(z)\\ &+ \int \chi(x,z)\big(L^{\nabla^{\pi^{-1}TB}}_{X^H}\mathrm{str}(G(x,z,z))\big)\mu_x(z)\\ &\qquad + \int \chi(x,z)\,\mathrm{str}(G(x,z,z))(L_{X^H}\mu_x(z)).\end{aligned}$$

Differentiating under the integral sign is valid because we knew a priori that the integrands are all L^1. Since $L_{X^H}\mu_x(z)$ equals $\mu_x(z)$ multiplied by some bounded functions, it follows that the last term $\int \chi(x,z)\,\mathrm{str}(G(x,z,z))(L_{X^H}\mu_x(z))$ is $O(t^{-\gamma'})$.

For the first term, we write $L_{X^H}\chi(x,z) = \sum_{g\in G}(g^*\chi)(x,z)(L_{X^H}\chi)(x,z)$. The sum is finite because $L_{X^H}\chi$ is compactly supported. By G-invariance,

$$\int (g^*\chi)(x,z)\,\mathrm{str}(G(x,z,z))\mu_x(z) = \int \chi(x,z)\,\mathrm{str}(G(x,z,z))\mu_x(z).$$

Since $(L_{X^H}\chi)(x,z)$ is bounded, it follows that $\int(L_{X^H}\chi)(x,z)\,\mathrm{str}(G(x,z,z))\mu_x(z)$ is also $O(t^{-\gamma'})$.

As for the second term, we differentiate under the integral sign, then use the Leibniz rule to get that there exists a constant $C_1 > 0$ such that

$$\begin{aligned}
&\big|L_{X^H}^{\nabla^{\pi^{-1}TB}}\,\mathrm{str}(G(x,z,z))\big| \\
&\quad\le C_1\bigg(\int_{Z_x}\big|L_{X^H}^{\nabla^{\wedge^\bullet H'\otimes\wedge^\bullet V'\otimes\hat{E}}} e^{-\eth(t/2)^2}(x,z,y)\big|\big|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}(x,y,z)\big|\mu_x(y) \\
&\qquad + \int_{Z_x}\big|e^{-\eth(t/2)^2}(x,z,y)\big|\big|L_{X^H}^{\nabla^{\wedge^\bullet H'\otimes\wedge^\bullet V'\otimes\hat{E}}}(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2})(x,y,z)\big|\mu_x(y) \\
&\qquad + \int_{Z_x}\big|e^{-\eth(t/2)^2}(x,z,y)\big|\big|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}(x,y,z)\big|\,\sup|L_{X^H}\mu|\mu_x(y)\bigg),
\end{aligned}$$

and

$$\begin{aligned}
&\int_B\bigg|\int_{Z_x}\chi(x,z)\big(L_{X^H}^{\nabla^{\pi^{-1}TB}}\,\mathrm{str}(G(x,z,z))\big)\mu_x(z)\bigg|^2\mu_B(x) \\
&\quad\le C_1\big(\|e^{-\eth(t/2)^2}\|^2_{\mathrm{HS}\,1}\|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\|^2_{\mathrm{HS}\,0} + \|e^{-\eth(t/2)^2}\|^2_{\mathrm{HS}\,0}\|e^{-\eth(t/2)^2} \\
&\qquad - \Pi_0 e^{(\Omega\Pi_0)^2}\|^2_{\mathrm{HS}\,1} + \sup|L_{X^H}\mu|\,\|e^{-\eth(t/2)^2}\|^2_{\mathrm{HS}\,1}\|e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2}\|^2_{\mathrm{HS}\,0}\big) \\
&\quad= O(t^{-\gamma'}).
\end{aligned}$$

Clearly the above arguments can be repeated and one concludes that all Sobolev norms of $\mathrm{str}_\Psi(G)$ are $O(t^{-\gamma'})$.

By exactly the same arguments, we have as $t \to \infty$,

$$\mathrm{str}_\Psi\big(2^{-1}N(e^{-\eth(t/2)^2} - \Pi_0 e^{(\Omega\Pi_0)^2})\Pi_0 e^{(\Omega\Pi_0)^2}\big) = O(t^{-\gamma'}),$$

in all Sobolev norms.

As for $\operatorname{str}_\Psi(2^{-1}N(D(t)^2e^{-\eth(t)^2}))$, one has $D(t)^2=2(2^{-N_\Omega/2}D\big(\frac{t}{2}\big)^2 2^{N_\Omega/2})$. Thus

$$\begin{aligned}\operatorname{str}_\Psi&\big(\tfrac{1}{2}N\big(D(t)^2e^{-\eth(t)^2}\big)\big)\\&=2^{-N_\Omega/2}\operatorname{str}_\Psi\big(N\big(D\big(\tfrac12 t\big)^2e^{-\eth(t/2)^2}e^{-\eth(t/2)^2}-\Pi_0(\Omega\Pi_0)^2e^{(\Omega\Pi_0)^2}e^{(\Omega\Pi_0)^2}\big)\big)\\&=2^{-N_\Omega/2}\operatorname{str}_\Psi\big(N\big(D\big(\tfrac12 t\big)^2e^{-\eth(t/2)^2}-\Pi_0(\Omega\Pi_0)^2e^{(\Omega\Pi_0)^2}\big)e^{-\eth(t/2)^2}\big)\\&\qquad\qquad-2^{-N_\Omega/2}\operatorname{str}_\Psi\big(N\Pi_0(\Omega\Pi_0)^2e^{(\Omega\Pi_0)^2}(e^{-\eth(t/2)^2}-e^{(\Omega\Pi_0)^2})\big),\end{aligned}$$

which is also $O(t^{-\gamma'})$ as $t\to\infty$ by similar arguments.

By the Sobolev embedding theorem (for the compact manifold B), it follows that

$$-F^\wedge(t)+\tfrac12\operatorname{str}_\Psi(N\Pi_0)+\big(\tfrac14\dim(Z)\operatorname{rk}(E)\operatorname{str}_\Psi(\Pi_0)-\tfrac12\operatorname{str}_\Psi(N\Pi_0)\big)(1-2t)e^{-t}$$

and all its derivatives are $O(t^{-\gamma'})$ uniformly.

Finally, since all derivatives of the t-integrand in Definition 4.5 are L^1, derivatives of τ exist and equal differentiations under the t-integration sign. Hence we conclude that the torsion τ is smooth. □

Remark 4.7. If Z is L^2-acyclic and of determinant class (see [Azzali et al. 2015, Def. 6.3]), the analogue of Remark 3.6 reads

$$\int_0^\infty\|e^{-t\Delta}\|^2_{\mathrm{HS}\,0}\,\frac{dt}{t}=\int_0^\infty\|e^{-t\Delta}\|_\tau\,\frac{dt}{t}<\infty$$

(note that $\Pi_0=0$ by hypothesis). Unlike having positive Novikov–Shubin invariant, the heat operator is not of determinant class in $\|\cdot\|_{\mathrm{HS}\,0}$.

Take a power series $f(x)=\sum a_jx^j$. For clarity, let h be the metric on $\wedge^\bullet V\otimes E$ and denote

$$f\big(\nabla^{\wedge^\bullet V'\otimes E},h\big):=\operatorname{str}\Big(\sum_j a_j\big(\tfrac12\big(\nabla^{\wedge^\bullet V'\otimes E}-(\nabla^{\wedge^\bullet V'\otimes E})^*\big)\big)^j\Big)\in\Gamma^\infty(\wedge^\bullet T^*M),$$

$$f\big(\nabla^{\wedge^\bullet V'\otimes E},h\big)_{H^\bullet(Z,E)}:=\operatorname{str}_\Psi\Big(\sum_j a_j\big(\tfrac12\Pi_0\big(\nabla^{\wedge^\bullet V'_\flat\otimes E_\flat}-(\nabla^{\wedge^\bullet V'_\flat\otimes E_\flat})^*\big)\Pi_0\big)^j\Big)\in\Gamma^\infty(\wedge^\bullet T^*B).$$

Note that the summations are only up to $\dim M$.

Let TZ be the vertical tangent bundle of the fiber bundle $M\to B$ and recall that we have chosen a splitting of TM and defined a Riemannian metric on TM. Let P^{TZ} denote the projection from TM to TZ. Let ∇^{TM} be the corresponding Levi-Civita connection on TM and define $\nabla^{TZ}=P^{TZ}\nabla^{TM}P^{TZ}$, a connection on TZ. The restriction of ∇^{TZ} to a fiber coincides with the Levi-Civita connection of the fiber. Let R^{TZ} be the curvature of ∇^{TZ}.

For N even, let $\mathrm{Pf} : \mathfrak{so}(N) \to \mathbb{R}$ denote the Pfaffian and put

$$e\left(TZ, \nabla^{TZ}\right) := \begin{cases} \mathrm{Pf}\left[\dfrac{R^{TZ}}{2\pi}\right] & \text{if } \dim(Z) \text{ is even,} \\ 0 & \text{if } \dim(Z) \text{ is odd.} \end{cases} \tag{23}$$

A classical argument [Bismut and Lott 1995; Ma and Zhang 2008; Azzali et al. 2015] then gives:

Corollary 4.8. *If* $\dim Z = 2n$ *is even one has the transgression formula*

$$d\tau(x) = \int_{Z_x} \chi(x, z) e(TZ, \nabla^{TZ}) f\left(\nabla^{\wedge^\bullet V' \otimes E}\right) - f\left(\nabla^{\wedge^\bullet V' \otimes E}\right)_{H^\bullet(Z,E)},$$

with $f(x) = xe^{x^2}$.

Now let h_l be a family of G-invariant metrics on $\wedge^\bullet V \otimes E$, $l \in [0, 1]$. Define

$$\tilde{f}\left(\nabla^{\wedge^\bullet V' \otimes E}, h_l\right) := \int_0^1 (2\pi\sqrt{-1})^{N_\Omega/2} \operatorname{str}\left((h_l)^{-1} \frac{dh_l}{dl} f'\left(\nabla^{\wedge^\bullet V' \otimes E}, h_l\right)\right) dl,$$

and similarly for $\tilde{f}\left(\nabla^{\wedge^\bullet V' \otimes E}, h_l\right)_{H^\bullet(Z,E)}$. Note that $f'\left(\nabla^{\wedge^\bullet V' \otimes E}, h_l\right)$ uses the adjoint connection with respect to h_l.

Let $\hat{e}\left(TZ, \nabla^{TZ,0}, \nabla^{TZ,1}\right) \in Q^M / Q^{M,0}$ (see [Bismut and Lott 1995]) be the secondary class associated to the Euler class. Its representatives are forms of degree $\dim(Z) - 1$ such that

$$d\hat{e}\left(TZ, \nabla^{TZ,0}, \nabla^{TZ,1}\right) = e\left(TZ, \nabla^{TZ,1}\right) - e\left(TZ, \nabla^{TZ,0}\right). \tag{24}$$

If $\dim(Z)$ is odd, we take $\hat{e}\left(TZ, \nabla^{TZ,0}, \nabla^{TZ,1}\right)$ to be zero.

One has an anomaly formula [Bismut and Lott 1995, Theorem 3.24].

Lemma 4.9. *Modulo exact forms*

$$\begin{aligned} \tau_1 - \tau_0 = &\int_{Z_x} \chi(x, z) \hat{e}(TZ, \nabla^{TZ,0}, \nabla^{TZ,1}) f\left(\nabla^{\wedge^\bullet V' \otimes E}, h_0\right) \\ &+ \int_{Z_x} \chi(x, z) e(TZ, \nabla^{TZ,1}) \tilde{f}\left(\nabla^{\wedge^\bullet V' \otimes E}, h_l\right) - \tilde{f}\left(\nabla^{\wedge^\bullet V'_\flat \otimes E_\flat}, h_l\right)_{H^\bullet(Z,E)}. \end{aligned} \tag{25}$$

In particular, the degree 0 part of equation (25) is the anomaly formula for the L^2-Ray–Singer analytic torsion, which is a special case of [Zhang 2005, Theorem 3.4].

Remark 4.10. Let $Z_0 \to M_0 \to B$ be a fiber bundle with compact fiber Z_0, $Z \to M \to B$ be the normal covering of the fiber bundle $Z_0 \to M_0 \to B$. Then one can define the Bismut–Lott and L^2-analytic torsion form $\tau_{M_0 \to B}$, $\tau_{M_0 \to B} \in \Gamma^\infty(\wedge^\bullet T^* B)$, and one has the respective transgression formulas

$$d\tau_{M_0\to B} = \int_{\pi_0^{-1}(x)} e(TZ_0, \nabla^{TZ_0}) f\big(\nabla^{\wedge^\bullet V_0'\otimes E_0}\big) - f\big(\nabla^{\wedge^\bullet V_0'\otimes E_0}\big)_{H^\bullet(Z_0,E_0)},$$

$$d\tau_{M\to B} = \int_{Z_x} \chi(x,z) e(TZ, \nabla^{TZ}) f\big(\nabla^{\wedge^\bullet V'\otimes E}\big) - f\big(\nabla^{\wedge^\bullet V'\otimes E}\big)_{H^\bullet(Z,E)}.$$

Suppose further that the de Rham cohomologies are trivial:

$$H^\bullet(Z_0, E|_{Z_0}) = H^\bullet_{L^2}(Z, E|_Z) = \{0\}.$$

Then $d(\tau_{M\to B} - \tau_{M_0\to B}) = 0$. Hence $\tau_{M\to B} - \tau_{M_0\to B}$ defines some class in the de Rham cohomology of B. We also remark that this form was also mentioned in [Azzali et al. 2015, Remark 7.5], as a weakly closed form.

Acknowledgements

The authors are very grateful to the referees for their very careful reading of the manuscript of the paper and many valuable suggestions. Su was supported by NSFC 11571183.

References

[Álvarez López and Kordyukov 2001] J. A. Álvarez López and Y. A. Kordyukov, "Long time behavior of leafwise heat flow for Riemannian foliations", *Compositio Math.* **125**:2 (2001), 129–153. MR Zbl

[Azzali et al. 2015] S. Azzali, S. Goette, and T. Schick, "Large time limit and local L^2-index theorems for families", *J. Noncommut. Geom.* **9**:2 (2015), 621–664. MR Zbl

[Berline et al. 1992] N. Berline, E. Getzler, and M. Vergne, *Heat kernels and Dirac operators*, Grundlehren der Mathematischen Wissenschaften **298**, Springer, 1992. MR Zbl

[Bismut 1986] J.-M. Bismut, "The Atiyah–Singer index theorem for families of Dirac operators: two heat equation proofs", *Invent. Math.* **83**:1 (1986), 91–151. MR Zbl

[Bismut and Lott 1995] J.-M. Bismut and J. Lott, "Flat vector bundles, direct images and higher real analytic torsion", *J. Amer. Math. Soc.* **8**:2 (1995), 291–363. 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.

[Carey and Mathai 1992] A. L. Carey and V. Mathai, "L^2-torsion invariants", *J. Funct. Anal.* **110**:2 (1992), 377–409. MR Zbl

[Gong and Rothenberg 1996] D. Gong and M. Rothenberg, "Analytic torsion forms on non-compact fiber bundles", preprint, 1996.

[Gorokhovsky and Lott 2006] A. Gorokhovsky and J. Lott, "Local index theory over foliation groupoids", *Adv. Math.* **204**:2 (2006), 413–447. MR Zbl

[Heitsch 1995] J. L. Heitsch, "Bismut superconnections and the Chern character for Dirac operators on foliated manifolds", *K-Theory* **9**:6 (1995), 507–528. MR Zbl

[Heitsch and Lazarov 2002] J. L. Heitsch and C. Lazarov, "Riemann–Roch–Grothendieck and torsion for foliations", *J. Geom. Anal.* **12**:3 (2002), 437–468. MR Zbl

[Lott 1992] J. Lott, “Heat kernels on covering spaces and topological invariants”, *J. Differential Geom.* **35**:2 (1992), 471–510. MR Zbl

[Ma and Zhang 2008] X. Ma and W. Zhang, “Eta-invariants, torsion forms and flat vector bundles”, *Math. Ann.* **340**:3 (2008), 569–624. MR Zbl

[Mathai 1992] V. Mathai, “L^2-analytic torsion”, *J. Funct. Anal.* **107**:2 (1992), 369–386. MR Zbl

[Milnor 1966] J. Milnor, “Whitehead torsion”, *Bull. Amer. Math. Soc.* **72** (1966), 358–426. MR Zbl

[Nistor et al. 1999] V. Nistor, A. Weinstein, and P. Xu, “Pseudodifferential operators on differential groupoids”, *Pacific J. Math.* **189**:1 (1999), 117–152. 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

[Shubin 1992] M. A. Shubin, “Spectral theory of elliptic operators on noncompact manifolds”, pp. 35–108 in *Méthodes semi-classiques, I* (Nantes, 1991), Astérisque **207**, 1992. MR Zbl

[Zhang 2005] W. Zhang, “An extended Cheeger–Müller theorem for covering spaces”, *Topology* **44**:6 (2005), 1093–1131. MR Zbl

Received June 9, 2014. Revised July 16, 2016.

BING KWAN SO
MATHEMATICS INSTITUTE
JILIN UNIVERSITY
2699 QIANJIN STREET
CHANGCHUN, 130012
CHINA
bkso@graduate.hku.hk

GUANGXIANG SU
CHERN INSTITUTE OF MATHEMATICS AND LPMC
NANKAI UNIVERSITY
TIANJIN, 300071
CHINA
guangxiangsu@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.183

THICK SUBCATEGORIES OVER ISOLATED SINGULARITIES

RYO TAKAHASHI

We study classifying thick subcategories of the category of finitely generated modules and its bounded derived category for a local ring with an isolated singularity.

1. Introduction

Let R be a commutative noetherian local ring. We denote by $\mathsf{mod}\,R$ the category of finitely generated R-modules, and by $\mathsf{D}^{\mathsf{b}}(R)$ the bounded derived category of $\mathsf{mod}\,R$.

First, we consider classifying thick subcategories of the abelian category $\mathsf{mod}\,R$. In general, thick subcategories are much more than Serre subcategories; even when R is a hypersurface, the cardinality of thick subcategories of $\mathsf{mod}\,R$ containing R is equal to that of specialization-closed subsets of the singular locus [Takahashi 2010; 2013b], while the only Serre subcategory of $\mathsf{mod}\,R$ containing R is the whole category $\mathsf{mod}\,R$.

We prove the following structure theorem of thick closures:

Theorem 1.1. *Let R be a local ring with residue field k, and suppose that R has an isolated singularity. For each nonzero finitely generated R-module M one has*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\} = \mathsf{thick}_{\mathsf{mod}\,R}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp} M\}$$

of thick closures, provided that one of the following three conditions is satisfied.

(i) *M is locally free on the punctured spectrum of R.*

(ii) *R has (Krull) dimension at most two.*

(iii) *R has prime characteristic and M is (not necessarily maximal) Cohen–Macaulay.*

As a byproduct of the above theorem and its proof, we obtain the following result. Denote by $\operatorname{Nesc}(R)$ the set of nonempty specialization-closed subsets of $\operatorname{Spec} R$.

The author was partly supported by JSPS Grant-in-Aid for Scientific Research 25400038, 16K05098.
MSC2010: 13C60, 13D09, 18E30.
Keywords: Cohen–Macaulay ring/module, derived category, isolated singularity, module category, specialization-closed subset, thick subcategory.

Theorem 1.2. (1) *Let R be a local ring with residue field k. Suppose that R has an isolated singularity and dimension at most two. Then taking the supports gives a one-to-one correspondence between the set of thick subcategories of* $\operatorname{mod} R$ *containing k and* $\operatorname{Nesc}(R)$. *In particular, all the thick subcategories of* $\operatorname{mod} R$ *containing k are Serre.*

(2) *Let R be a regular local ring of positive characteristic. Then the thick closure in* $\operatorname{mod} R$ *of each nonzero R-module of finite length consists of all R-modules of finite length.*

Next, we consider classifying thick subcategories of the triangulated category $\mathsf{D}^{\mathsf{b}}(R)$. Stevenson [2014] completely classified the thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$ in the case where R is a complete intersection. Thus, our next goal is to classify the thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$ for a non-complete-intersection local ring R. However, this problem itself turns out to be quite hard, and it would be a reasonable approach to consider classifying the thick subcategories satisfying a certain condition which all the thick subcategories satisfy over complete intersections. The standard and costandard conditions are such ones; a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ is called *standard* (resp. *costandard*) if it contains a nonzero object of finite projective (resp. injective) dimension. Dwyer, Greenlees and Iyengar [Dwyer et al. 2006] showed that if R is a complete intersection, then every nonzero thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ is standard and costandard. We show the following classification theorem of standard and costandard thick subcategories:

Theorem 1.3. *Let R be a singular Cohen–Macaulay local ring with an isolated singularity. Assume that R is complete and has infinite residue field.*

(1) *If R is a hypersurface, then there is a one-to-one correspondence between the set of nonzero thick subcategories of* $\mathsf{D}^{\mathsf{b}}(R)$ *and the disjoint union of two copies of* $\operatorname{Nesc}(R)$.

(2) *If R has minimal multiplicity, then there is a one-to-one correspondence between the set of standard thick subcategories of* $\mathsf{D}^{\mathsf{b}}(R)$ *and the disjoint union of two copies of* $\operatorname{Nesc}(R)$.

(3) *If either R is non-Gorenstein and almost Gorenstein or R is of finite CM-representation type, then taking the supports gives a one-to-one correspondence between the set of standard and costandard thick subcategories of* $\mathsf{D}^{\mathsf{b}}(R)$ *and* $\operatorname{Nesc}(R)$.

In fact, the bijections in the first and second assertions are also explicitly described. The first assertion can also be deduced from [Stevenson 2014].

This paper is organized as follows. Section 2 is for preliminaries. The proof of Theorem 1.1 is divided into Sections 3, 4 and 5. In Section 6 we classify the thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$ containing k. Applications of this, including Theorem 1.3, are given in Sections 7, 8 and 9.

2. Fundamental definitions

Throughout this paper, let R be a commutative noetherian ring. We assume that all modules are finitely generated, and that all subcategories are nonempty, full and closed under isomorphism. Denote by $\operatorname{mod} R$ the category of (finitely generated) R-modules, by $\mathsf{C}^{\mathsf{b}}(R)$ the category of bounded complexes of (finitely generated) R-modules and by $\mathsf{D}^{\mathsf{b}}(R)$ the bounded derived category of $\operatorname{mod} R$. Note that $\operatorname{mod} R$ and $\mathsf{C}^{\mathsf{b}}(R)$ are abelian, while $\mathsf{D}^{\mathsf{b}}(R)$ is triangulated.

Definition 2.1. (1) A subcategory $\mathcal{X}$ of $\operatorname{mod} R$ is called *Serre* if it is closed under submodules, quotient modules and extensions.

(2) A subcategory $\mathcal{X}$ of $\operatorname{mod} R$ (resp. $\mathsf{C}^{\mathsf{b}}(R)$, $\mathsf{D}^{\mathsf{b}}(R)$) is called *thick* if it is closed under direct summands and satisfies the 2-out-of-3 property for short exact sequences of modules (resp. short exact sequences of complexes and closed under shifts, exact triangles).

(3) A subset S of $\operatorname{Spec} R$ is called *specialization-closed* if S contains $\mathrm{V}(\mathfrak{p})$ for all $\mathfrak{p} \in S$. Note that this is equivalent to saying that S is a union of closed subsets of $\operatorname{Spec} R$.

(4) (a) For each $M \in \operatorname{mod} R$ we denote by $\operatorname{Supp}_R M$ the set of prime ideals $\mathfrak{p}$ of R with $M_{\mathfrak{p}} \ncong 0$ in $\operatorname{mod} R_{\mathfrak{p}}$, and call this the *support* of M in $\operatorname{mod} R$. This is a closed subset of $\operatorname{Spec} R$.

(b) The *support* of a subcategory $\mathcal{X}$ of $\operatorname{mod} R$ is defined by $\operatorname{Supp}_R \mathcal{X} = \bigcup_{X \in \mathcal{X}} \operatorname{Supp}_R X$. This is a specialization-closed subset of $\operatorname{Spec} R$.

(c) For a subset S of $\operatorname{Spec} R$ we denote by $\operatorname{Supp}^{-1}_{\operatorname{mod} R} S$ the subcategory of $\operatorname{mod} R$ consisting of all modules whose supports are contained in S. This is a Serre subcategory of $\operatorname{mod} R$.

(d) The *support* of an object $X \in \mathsf{D}^{\mathsf{b}}(R)$, denoted by $\operatorname{Supp}_R X$, is defined as the support of its homology $\mathrm{H}(X)$. Hence this is a closed subset.

(e) The *support* of a subcategory $\mathcal{X}$ of $\mathsf{D}^{\mathsf{b}}(R)$ is defined by $\operatorname{Supp}_R \mathcal{X} = \bigcup_{X \in \mathcal{X}} \operatorname{Supp}_R X$. This is a specialization-closed subset of $\operatorname{Spec} R$.

(f) For a subset S of $\operatorname{Spec} R$ we denote by $\operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)} S$ the subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ consisting of objects whose supports are contained in S. This is a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$.

(5) A *perfect* complex is by definition (a complex quasi-isomorphic to) a bounded complex of finitely generated projective modules. We denote by $\mathsf{D}_{\mathsf{perf}}(R)$ the subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ consisting of perfect complexes. This is a thick

subcategory of $\mathsf{D^b}(R)$, and hence a triangulated category. For each subset S of $\operatorname{Spec} R$ we set $\operatorname{Supp}^{-1}_{\mathsf{D_{perf}}(R)} S = (\operatorname{Supp}^{-1}_{\mathsf{D^b}(R)} S) \cap \mathsf{D_{perf}}(R)$.

(6) Let $(R, \mathfrak{m})$ be a local ring, and let M be an R-module. Choose a minimal free resolution $F = (\cdots \to F_n \xrightarrow{\partial_n} F_{n-1} \to \cdots \to F_1 \xrightarrow{\partial_1} F_0 \to 0)$ of M. We define the *n-th syzygy* $\Omega^n_R M$ and the *transpose* $\operatorname{Tr}_R M$ of M by $\Omega^n_R M = \operatorname{Im}(\partial_n)$ and $\operatorname{Tr}_R M = \operatorname{Cok}(\partial_1^*)$, where we set $(-)^* = \operatorname{Hom}_R(-, R)$. One has $\Omega^n_R M \subseteq \mathfrak{m}F_{n-1}$ and $M^* \cong \Omega^2_R \operatorname{Tr}_R M \oplus R^{\oplus t}$ for some $t \geq 0$.

Since $\mathsf{C^b}(R)$ is abelian, we can define a complex over $\mathsf{C^b}(R)$. More precisely, a *complex* of objects of $\mathsf{C^b}(R)$ is a sequence $X = (\cdots \xrightarrow{d_{i+1}} X_i \xrightarrow{d_i} X_{i-1} \xrightarrow{d_{i-1}} \cdots)$ of morphisms $d_i : X_i \to X_{i-1}$ in $\mathsf{C^b}(R)$ with $d_i d_{i+1} = 0$. We introduce a Koszul complex on a complex of R-modules.

Definition 2.2. Let X be an object of $\mathsf{C^b}(R)$. Let $\boldsymbol{x} = x_1, \dots, x_n$ be a sequence of elements of R, and let $K = \mathrm{K}(\boldsymbol{x}, R) = (0 \to K_n \xrightarrow{\partial_n} K_{n-1} \to \cdots \to K_1 \xrightarrow{\partial_1} K_0 \to 0)$ be the Koszul complex of $\boldsymbol{x}$ on R. We define the *Koszul complex* $\mathrm{K}(\boldsymbol{x}, X)$ of $\boldsymbol{x}$ on X by

$$\mathrm{K}(\boldsymbol{x},X) = (0 \to K_n \otimes_R X \xrightarrow{\partial_n \otimes_R X} K_{n-1} \otimes_R X \to \cdots \to K_1 \otimes_R X \xrightarrow{\partial_1 \otimes_R X} K_0 \otimes_R X \to 0),$$

where each $\partial_i \otimes_R X$ is a usual chain map, that is, a morphism in $\mathsf{C^b}(R)$. The Koszul complex $\mathrm{K}(\boldsymbol{x}, X)$ is a complex of objects of $\mathsf{C^b}(R)$.

Let $\mathcal{C}$ be one of the categories $\operatorname{mod} R$, $\mathsf{C^b}(R)$ and $\mathsf{D^b}(R)$. For a subcategory $\mathcal{M}$ of $\mathcal{C}$, the *thick closure* of $\mathcal{M}$ in $\mathcal{C}$, denoted by $\operatorname{thick}_{\mathcal{C}} \mathcal{M}$, is by definition the smallest thick subcategory of $\mathcal{C}$ containing $\mathcal{M}$. The proof of the following lemma is standard and omitted.

Lemma 2.3. (1) *Let $\mathcal{C}$ be one of the categories* $\operatorname{mod} R$, $\mathsf{C^b}(R)$ *and* $\mathsf{D^b}(R)$. *Let $\mathcal{X}$ be a thick subcategory of $\mathcal{C}$. Let $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = M$ be a filtration in* $\operatorname{mod} R$. *If M_i/M_{i-1} is in $\mathcal{X}$ for each $1 \leq i \leq n$, then so is M. In particular, M is in* $\operatorname{thick}_{\mathcal{C}}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R M\}$.

(2) *Let $\mathcal{X}$ be a thick subcategory of* $\operatorname{mod} R$, *and $X = (0 \to X_n \to \cdots \to X_1 \to X_0 \to 0)$ be a complex of R-modules in $\mathcal{X}$. If $\mathrm{H}_i(X) \in \mathcal{X}$ for all $1 \leq i \leq n$, then $\mathrm{H}_0(X) \in \mathcal{X}$.*

(3) *Let $\mathcal{X}$ be a thick subcategory of* $\mathsf{D^b}(R)$. *Let $C \in \mathsf{D^b}(R)$. If $\mathrm{H}(C)$ is in $\mathcal{X}$, then so is C.*

(4) *Let $X = (0 \to X^s \to X^{s+1} \to \cdots \to X^t \to 0)$ be a complex of R-modules. Then X belongs to* $\operatorname{thick}_{\mathsf{C^b}(R)}\{X^s, X^{s+1}, \dots, X^t\}$.

(5) *Let $\mathcal{X}$ be a thick subcategory of* $\mathsf{C^b}(R)$. *Let $X = (0 \to X_n \to \cdots \to X_0 \to 0)$ be a complex of objects of $\mathsf{C^b}(R)$ with $X_0, \dots, X_n \in \mathcal{X}$. If $\mathrm{H}_i(X) \in \mathcal{X}$ for all $1 \leq i \leq n$, then $\mathrm{H}_0(X) \in \mathcal{X}$.*

3. Modules locally free on the punctured spectrum

Let R be a local ring with residue field k. In this section, we study the structure of the thick closure of k and M in $\mathsf{mod}\, R$ when M is locally free on the punctured spectrum of R.

Lemma 3.1. *Let $\mathfrak{p}$ be a prime ideal of R.*

(1) *Suppose that $R_\mathfrak{p}$ is a regular local ring of dimension n. Then for each $0 \le i \le n$ there is an ideal $J = (x_1, \dots, x_i) \subseteq \mathfrak{p}$ with $\operatorname{ht} J = i$ such that $R_\mathfrak{p}/JR_\mathfrak{p}$ is a regular local ring of dimension $n - i$. In particular, there is an ideal $I = (x_1, \dots, x_n)$ of height n with $IR_\mathfrak{p} = \mathfrak{p}R_\mathfrak{p}$.*

(2) *Let I be an ideal of R with $IR_\mathfrak{p} = \mathfrak{p}R_\mathfrak{p}$. Then there exists an exact sequence $0 \to R/I \to R/\mathfrak{p} \oplus R/\mathfrak{q} \to R/J \to 0$ of R-modules such that J strictly contains $\mathfrak{p}$.*

Proof. (1) We use induction on n. First of all, note that the assertion evidently holds for $i = 0$. When $n = 0$, we have $i = 0$, and we are done. Let $n \ge 1$. We may assume $1 \le i \le n$, so $0 \le i - 1 \le n - 1$. The induction hypothesis implies that there is an ideal $K = (x_1, \dots, x_{i-1}) \subseteq \mathfrak{p}$ with $\operatorname{ht} K = i - 1$ such that $R_\mathfrak{p}/KR_\mathfrak{p}$ is a regular local ring of dimension $n - i + 1$. Set $\bar{R} = R/K$ and $\bar{\mathfrak{p}} = \mathfrak{p}/K$. The local ring $\bar{R}_{\bar{\mathfrak{p}}}$ is regular and $\operatorname{ht} \bar{\mathfrak{p}} = \dim \bar{R}_{\bar{\mathfrak{p}}} = n - i + 1 > 0$. Nakayama's lemma shows that the symbolic power $\bar{\mathfrak{p}}^{(2)} = \bar{\mathfrak{p}}^2 \bar{R}_{\bar{\mathfrak{p}}} \cap \bar{R}_{\bar{\mathfrak{p}}}$ is strictly contained in $\bar{\mathfrak{p}}$. By prime avoidance we find an element $\overline{x_i} \in \bar{\mathfrak{p}}$ that is not contained in the union of ideals in $\operatorname{Min} \bar{R} \cup \{\bar{\mathfrak{p}}^{(2)}\}$. It is easy to see that the ideal $J := K + (x_i)$ has height i and $R_\mathfrak{p}/JR_\mathfrak{p} = \bar{R}_{\bar{\mathfrak{p}}}/\overline{x_i}\, \bar{R}_{\bar{\mathfrak{p}}}$ is a regular local ring of dimension $n - i$.

(2) Since $\mathfrak{p}$ is a minimal prime of I and $\mathfrak{p} = IR_\mathfrak{p} \cap R$, we see that $\mathfrak{p}$ is a $\mathfrak{p}$-primary component of I. Hence we can write $I = \mathfrak{p} \cap \mathfrak{q}$ for some ideal $\mathfrak{q}$ of R that is not contained in $\mathfrak{p}$ (when I is itself $\mathfrak{p}$-primary, we can take $\mathfrak{q} = R$). There is an exact sequence $0 \to R/I \to R/\mathfrak{p} \oplus R/\mathfrak{q} \to R/J \to 0$, where $J := \mathfrak{p} + \mathfrak{q}$ strictly contains $\mathfrak{p}$. $\square$

The following result plays a key role in the proof of the main result of this section:

Lemma 3.2. *Suppose that R is locally Cohen–Macaulay on the punctured spectrum. Let $\boldsymbol{x} = x_1, \dots, x_n$ be a sequence of elements of R generating an ideal of height n. Let M be an R-module locally free on the punctured spectrum of R. Then for each $i > 0$ the i-th Koszul homology $\mathrm{H}_i(\boldsymbol{x}, M)$ has finite length as an R-module.*

Proof. Pick any nonmaximal prime ideal $\mathfrak{p}$ of R. We want to show that $\mathrm{H}_i(\boldsymbol{x}, M)_\mathfrak{p}$ vanishes for all $i > 0$. This $R_\mathfrak{p}$-module is isomorphic to $\mathrm{H}_i(\boldsymbol{x}, M_\mathfrak{p})$, and $M_\mathfrak{p}$ is a free $R_\mathfrak{p}$-module. Hence it suffices to show that $\mathrm{H}_i(\boldsymbol{x}, R_\mathfrak{p}) = 0$ for all $i > 0$. This holds true if $\mathfrak{p}$ does not contain $\boldsymbol{x}$, since $\boldsymbol{x}\, \mathrm{H}_i(\boldsymbol{x}, R_\mathfrak{p}) = 0$ for all $i \in \mathbb{Z}$. Let us consider the case where $\mathfrak{p}$ contains $\boldsymbol{x}$. We then have

$$n \geq \operatorname{ht}(\boldsymbol{x}R_{\mathfrak{p}}) = \inf\{\operatorname{ht} Q \mid Q \in \mathrm{V}(\boldsymbol{x}R_{\mathfrak{p}})\} = \inf\{\operatorname{ht}\mathfrak{q} \mid \mathfrak{q} \in \mathrm{V}(\boldsymbol{x}R),\ \mathfrak{q} \subseteq \mathfrak{p}\}$$
$$\geq \inf\{\operatorname{ht}\mathfrak{q} \mid \mathfrak{q} \in \mathrm{V}(\boldsymbol{x}R)\} = \operatorname{ht}(\boldsymbol{x}R) = n,$$

where the first inequality follows from Krull's height theorem. Hence the ideal $\boldsymbol{x}R_{\mathfrak{p}}$ generated by n elements has height n. Since $R_{\mathfrak{p}}$ is a Cohen–Macaulay local ring by assumption, $\boldsymbol{x}$ is an $R_{\mathfrak{p}}$-sequence. Therefore $\mathrm{H}_i(\boldsymbol{x}, R_{\mathfrak{p}}) = 0$ for all $i > 0$. □

Recall that a local ring R is said to have an *isolated singularity* if for every nonmaximal prime ideal $\mathfrak{p}$ of R the local ring $R_{\mathfrak{p}}$ is regular. The following is the main result of this section.

Theorem 3.3. *Let $(R, \mathfrak{m}, k)$ be a local ring with an isolated singularity. Let M be a nonzero R-module which is locally free on the punctured spectrum of R. Then*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\} = \mathsf{thick}_{\mathsf{mod}\,R}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R M\}.$$

Proof. As M is nonzero, the maximal ideal $\mathfrak{m}$ is in the support of M. Lemma 2.3(1) implies that the inclusion ($\subseteq$) holds. We show the opposite inclusion ($\supseteq$). Set $\mathcal{X} = \mathsf{thick}\{k, M\}$. The proof will be completed once we prove that $R/I \in \mathcal{X}$ for all ideals I of R with $\mathrm{V}(I) \subseteq \operatorname{Supp} M$. Suppose that this does not hold; we will be done if we derive a contradiction. The set of ideals

$$\{I \subseteq R \mid R/I \notin \mathcal{X},\ \mathrm{V}(I) \subseteq \operatorname{Supp} M\}$$

is nonempty, and this has a maximal element P with respect to the inclusion relation, as R is noetherian. We establish a claim.

Claim. One has $\mathfrak{m} \neq P \in \operatorname{Supp} M$. Every R-module L with $\operatorname{Supp} L \subseteq \mathrm{V}(P) - \{P\}$ is in $\mathcal{X}$.

Proof of Claim. Since P is in the above set of ideals, the module R/P is not in $\mathcal{X}$ and $\mathrm{V}(P)$ is contained in $\operatorname{Supp} M$. As k is in $\mathcal{X}$, we have $P \neq \mathfrak{m}$. It remains to show that P is a prime ideal. Take a filtration $0 = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_n = R/P$ such that each N_i/N_{i-1} is isomorphic to $R/\mathfrak{p}_i$ for some prime ideal $\mathfrak{p}_i$ in $\operatorname{Supp}_R(R/P) = \mathrm{V}(P)$. Assume that P is not a prime ideal. Then each $\mathfrak{p}_i$ strictly contains P, and the maximality of P implies $R/\mathfrak{p}_i \in \mathcal{X}$ for all $1 \leq i \leq n$. By Lemma 2.3(1) we have $R/P \in \mathcal{X}$. This contradiction shows that P is a prime ideal of R.

Take a filtration $0 = L_0 \subsetneq L_1 \subsetneq \cdots \subsetneq L_\ell = L$ such that for each i one has $L_i/L_{i-1} \cong R/\mathfrak{p}_i$ with $\mathfrak{p}_i \in \operatorname{Supp} L \subseteq \mathrm{V}(P) - \{P\}$. The $\mathfrak{p}_i$ strictly contain P, and the maximality of P implies $R/\mathfrak{p}_i \in \mathcal{X}$, which forces L to be in $\mathcal{X}$ by Lemma 2.3(1). □

Since P is a nonmaximal prime ideal by the Claim and R is an isolated singularity, the localization R_P is a regular local ring. By Lemma 3.1 there is an exact sequence

$$0 \to R/(\boldsymbol{x}) \to R/P \oplus R/Q \to R/J \to 0,$$

where $\boldsymbol{x} = x_1, \dots, x_n$ is a sequence of elements of R with $\mathrm{ht}(\boldsymbol{x}) = n$, and J strictly contains P. Applying the functor $- \otimes_R M$ to this gives rise to an exact sequence

$$\mathrm{Tor}_1^R(R/J, M) \xrightarrow{f} M/\boldsymbol{x}M \to M/PM \oplus M/QM \to M/JM \to 0.$$

The supports of M/JM and $\mathrm{Tor}_1^R(R/J, M)$ are contained in $\mathrm{V}(J)$, and so is the image C of the map f. As $\mathrm{V}(J)$ is contained in $\mathrm{V}(P) - \{P\}$, the Claim implies that M/JM and C are in $\mathcal{X}$.

Now, assume that M is locally free on the punctured spectrum. Then by Lemma 3.2 for each $i > 0$ the i-th Koszul homology $\mathrm{H}_i(\boldsymbol{x}, M)$ has finite length, and it is in $\mathcal{X}$. Each component of the Koszul complex $\mathrm{K}(\boldsymbol{x}, M)$ is a direct sum of copies of M, which is in $\mathcal{X}$. Lemma 2.3(2) implies $M/\boldsymbol{x}M = \mathrm{H}_0(\boldsymbol{x}, M) \in \mathcal{X}$. The induced exact sequence

$$0 \to C \to M/\boldsymbol{x}M \to M/PM \oplus M/QM \to M/JM \to 0$$

shows that M/PM is also in $\mathcal{X}$. As M/PM is a module over the domain R/P, it has a rank, say r. There is an exact sequence

$$0 \to (R/P)^{\oplus r} \to M/PM \to E \to 0$$

of R/P-modules with $\dim E < \dim R/P$. Since P is in the support of M by the Claim, Nakayama's lemma implies that it is also in the support of M/PM, and hence $r > 0$. It is easy to see that $\mathrm{Supp}_R E$ is contained in $\mathrm{V}(P) - \{P\}$, and the Claim implies $E \in \mathcal{X}$. As $M/PM \in \mathcal{X}$ and $r > 0$, the module R/P is in $\mathcal{X}$. This contradiction completes the proof of the theorem. □

Remark 3.4. We should remark that the equality in Theorem 3.3 is no longer true if we remove k from the left-hand side. The equality

$$\mathsf{thick}_{\mathsf{mod}\,R} M = \mathsf{thick}_{\mathsf{mod}\,R}\{R/\mathfrak{p} \mid \mathfrak{p} \in \mathrm{Supp}_R M\}$$

holds for $M = R$ if and only if R is regular. This is one of the reasons why we consider thick subcategories containing k. See also Remark 6.4 stated later.

Applying Theorem 3.3 to $M = R$ and using Lemma 2.3(1), we obtain the following. This is a special case of [Schoutens 2003, Theorem VI.8] and [Krause and Stevenson 2013, Proposition 9], and includes [Takahashi 2010, Corollary 2.7].

Corollary 3.5. *If $(R, \mathfrak{m}, k)$ is an isolated singularity, then*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, R\} = \mathsf{mod}\,R.$$

4. Rings of dimension at most two

In this section, we deal with the same problem as in the previous section for local rings with dimension at most 2.

Lemma 4.1. *Let $(R, \mathfrak{m})$ be local. Let M, N be R-modules, and $\mathfrak{p}$ a prime ideal. Assume $M_\mathfrak{p} \cong N_\mathfrak{p}$ and $M_\mathfrak{q} = 0 = N_\mathfrak{q}$ for all $\mathfrak{q} \in \operatorname{Spec} R - \{\mathfrak{p}, \mathfrak{m}\}$. Then*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\} = \mathsf{thick}_{\mathsf{mod}\,R}\{k, N\}.$$

Proof. Since $M_\mathfrak{p}$ is isomorphic to $N_\mathfrak{p}$, there is an exact sequence

$$0 \to K \to M \to N \to C \to 0$$

such that $K_\mathfrak{p} = 0 = C_\mathfrak{p}$. We have $K_\mathfrak{q} = 0 = C_\mathfrak{q}$ for all $\mathfrak{q} \in \operatorname{Spec} R - \{\mathfrak{p}, \mathfrak{m}\}$, so K, C have finite length. Hence they are in both $\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\}$ and $\mathsf{thick}_{\mathsf{mod}\,R}\{k, N\}$, and it is seen that $N \in \mathsf{thick}_{\mathsf{mod}\,R}\{k, M\}$ and $M \in \mathsf{thick}_{\mathsf{mod}\,R}\{k, N\}$. Thus the assertion follows. □

The next lemma is well known and also easy to prove, so we omit the proof.

Lemma 4.2. (1) *Let $x \in R$ be a nonzerodivisor, and let $n > 0$ be an integer. Then there exists a short exact sequence*

$$0 \to R/(x^n) \to R/(x^{n+1}) \oplus R/(x^{n-1}) \to R/(x^n) \to 0,$$

where $x^0 := 1$

(2) *Let S be a multiplicatively closed subset of R. Let $\sigma : 0 \to M_S \to X \to N_S \to 0$ be an exact sequence of R_S-modules. Then there exists an exact sequence $\tau : 0 \to M \to Y \to N \to 0$ of R-modules such that $X \cong Y_S$.*

For a module M over a local ring R we denote by $\operatorname{Assh} M$ the set of prime ideals $\mathfrak{p}$ in the support of M with $\dim R/\mathfrak{p} = \dim M$. The following is a similar type of result to Theorem 3.3.

Theorem 4.3. *Let $(R, \mathfrak{m}, k)$ be a local ring with an isolated singularity. Suppose that R has Krull dimension at most 2. Then for any nonzero R-module M one has the equality*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\} = \mathsf{thick}_{\mathsf{mod}\,R}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R M\}.$$

Proof. The inclusion ($\subseteq$) follows from Lemma 2.3(1) and the fact that $\mathfrak{m}$ supports M, so we prove the opposite inclusion ($\supseteq$). We may assume that R, M have positive (Krull) dimension.

(1) If $\dim R = 1$, then the assumption that R has an isolated singularity forces M to be locally free on the punctured spectrum, and Theorem 3.3 shows the assertion.

(2) If $\dim R = 2$, then M has dimension either 1 or 2. Taking the $\mathfrak{m}$-torsion submodule of M, we see that there is an exact sequence $0 \to L \to M \to N \to 0$ such that L has finite length and N is a nonzero module of positive depth. We have $\operatorname{Supp} M = \operatorname{Supp} N$ and $\mathsf{thick}\{k, M\} = \mathsf{thick}\{k, N\}$. Replacing M with N, we may assume that M has positive depth.

(a) Suppose $\dim M = 1$. Then M is a 1-dimensional Cohen–Macaulay module, and it follows from [Bruns and Herzog 1998, Theorem 2.1.2(a)] that one has $\operatorname{Ass} M = \operatorname{Min} M = \operatorname{Assh} M = \{\mathfrak{p}_1, \dots, \mathfrak{p}_n\}$, where each prime ideal $\mathfrak{p}_i$ is such that $\dim R/\mathfrak{p}_i = 1$.

Let $0 = M_1 \cap \cdots \cap M_n$ be an irredundant primary decomposition of the zero submodule 0 of M such that M_i is $\mathfrak{p}_i$-primary for $1 \le i \le n$. There are exact sequences

$$0 \to M/N_{i-1} \to M/M_i \oplus M/N_i \to M/M_i + N_i \to 0 \quad (1 \le i \le n-1),$$

where $N_i := M_{i+1} \cap \cdots \cap M_n$. Each $M/M_i + N_i$ has finite length, and we get $\mathsf{thick}\{k, M\} = \mathsf{thick}\{k, M/M_1, \dots, M/M_n\}$. For each $1 \le i \le n$ we have $\operatorname{Ass} M/M_i = \{\mathfrak{p}_i\}$, which especially says that M/M_i is a 1-dimensional Cohen–Macaulay R-module whose support contains $\mathfrak{p}_i$.

Fix a prime ideal $\mathfrak{p}$ in the support of M. We want to show that $R/\mathfrak{p}$ is in $\mathsf{thick}\{k, M\}$. For this, we may assume $\mathfrak{p} \ne \mathfrak{m}$, and then we have $\mathfrak{p} = \mathfrak{p}_\ell$ for some $1 \le \ell \le n$. Replacing M with M/M_ℓ, we may assume $\operatorname{Ass} M = \{\mathfrak{p}\}$ and $\dim R/\mathfrak{p} = 1$. Then $M_\mathfrak{p}$ is a nonzero $R_\mathfrak{p}$-module of finite length and $\operatorname{Supp} M = \{\mathfrak{p}, \mathfrak{m}\}$. As $R_\mathfrak{p}$ is either a field or a discrete valuation ring, the structure theorem of finitely generated modules over principal ideal domains implies that

$$M_\mathfrak{p} \cong (R_\mathfrak{p}/\mathfrak{p}^{a_1} R_\mathfrak{p})^{\oplus b_1} \oplus \cdots \oplus (R_\mathfrak{p}/\mathfrak{p}^{a_t} R_\mathfrak{p})^{\oplus b_t}$$

for some $t > 0$, $a_1 > \cdots > a_t > 0$ and $b_1, \dots, b_t > 0$. Setting

$$E = (R/\mathfrak{p}^{a_1})^{\oplus b_1} \oplus \cdots \oplus (R/\mathfrak{p}^{a_t})^{\oplus b_t},$$

we have $M_\mathfrak{p} \cong E_\mathfrak{p}$ and $\operatorname{Supp} E = \{\mathfrak{p}, \mathfrak{m}\}$. Lemma 4.1 implies $\mathsf{thick}\{k, M\} = \mathsf{thick}\{k, E\}$.

We claim $R/\mathfrak{p}^n \in \mathsf{thick}_{\mathsf{mod}\, R}\{k, R/\mathfrak{p}^{n+1}\}$ for all $n > 0$. In fact, there is an exact sequence of $R_\mathfrak{p}$-modules

$$0 \to R_\mathfrak{p}/\mathfrak{p}^{n+1} R_\mathfrak{p} \to (R_\mathfrak{p}/\mathfrak{p}^n R_\mathfrak{p}) \oplus (R_\mathfrak{p}/\mathfrak{p}^{n+2} R_\mathfrak{p}) \to R_\mathfrak{p}/\mathfrak{p}^{n+1} R_\mathfrak{p} \to 0;$$

this is trivial when $R_\mathfrak{p}$ is a field, and follows from Lemma 4.2(1) when $R_\mathfrak{p}$ is a discrete valuation ring. Put $V = R/\mathfrak{p}^n \oplus R/\mathfrak{p}^{n+2}$. Lemma 4.2(2) yields an exact sequence $0 \to R/\mathfrak{p}^{n+1} \to W \to R/\mathfrak{p}^{n+1} \to 0$ such that $V_\mathfrak{p} \cong W_\mathfrak{p}$. As $\operatorname{Supp} V = \operatorname{Supp} W = \{\mathfrak{p}, \mathfrak{m}\}$, Lemma 4.1 implies $\mathsf{thick}\{k, V\} = \mathsf{thick}\{k, W\}$. The claim follows.

Using the claim repeatedly, we observe that $R/\mathfrak{p}$ belongs to $\mathsf{thick}\{k, R/\mathfrak{p}^n\}$ for all $n > 0$. Hence $R/\mathfrak{p}$ is in $\mathsf{thick}\{k, E\}$, and therefore it is in $\mathsf{thick}\{k, M\}$, as desired.

(b) Suppose $\dim M = 2$. Set $(-)^* = \operatorname{Hom}_R(-, R)$, and let $\lambda : M \to M^{**}$ be the natural homomorphism. Extend this to the exact sequence

$$0 \to K \to M \xrightarrow{\lambda} M^{**} \to C \to 0.$$

The module M^{**} is a second syzygy, and we have $K \cong \mathrm{Ext}^1_R(\mathrm{Tr}\, M, R)$ and $C \cong \mathrm{Ext}^2_R(\mathrm{Tr}\, M, R)$ by [Auslander and Bridger 1969, Proposition (2.6)]. As R is a 2-dimensional isolated singularity, M^{**} is locally free on the punctured spectrum, K has dimension at most 1 and C has finite length. The image E of λ is nonzero and locally free on the punctured spectrum. Applying Theorem 3.3 to E yields

$$\mathsf{thick}\{k, E\} = \mathsf{thick}\{R/\mathfrak{p} \mid \mathfrak{p} \in \mathrm{Supp}\, E\}. \tag{4.3.1}$$

The above exact sequence induces a short exact sequence $\sigma : 0 \to K \to M \to E \to 0$. Hence

$$\mathrm{Supp}\, M = \mathrm{Supp}\, K \cup \mathrm{Supp}\, E. \tag{4.3.2}$$

Since M has positive depth, K is a Cohen–Macaulay R-module of dimension 1. By part **(a)** on the previous page we get

$$\mathsf{thick}\{k, K\} = \mathsf{thick}\{R/\mathfrak{p} \mid \mathfrak{p} \in \mathrm{Supp}\, K\}. \tag{4.3.3}$$

As E is locally free on the punctured spectrum, the R-module $\mathrm{Ext}^1_R(E, K)$ has finite length, and hence the annihilator $\mathfrak{a} = \mathrm{Ann}_R\, \mathrm{Ext}^1_R(E, K)$ is $\mathfrak{m}$-primary. Thus one can choose a K-regular element x in $\mathfrak{a}$. The choice of x implies that the exact sequence $x\sigma$ splits, and we observe that there is an exact sequence

$$0 \to M \to K \oplus E \to K/xK \to 0.$$

As K/xK has finite length, K and E belong to $\mathsf{thick}\{k, M\}$. The exact sequence σ implies that M is in $\mathsf{thick}\{K, E\}$, and hence

$$\mathsf{thick}\{k, M\} = \mathsf{thick}\{k, K, E\}. \tag{4.3.4}$$

Combining (4.3.1)–(4.3.4) implies $\mathsf{thick}\{k, M\} = \mathsf{thick}\{R/\mathfrak{p} \mid \mathfrak{p} \in \mathrm{Supp}\, M\}$. □

Corollary 4.4. *Let* $(R, \mathfrak{m}, k)$ *be a local ring with* $\dim R \le 2$ *and having an isolated singularity.*

(1) *If* $\mathcal{X}$ *is a thick subcategory of* $\mathsf{mod}\, R$ *containing* k, *then*

$$\mathrm{Supp}_R\, \mathcal{X} = \{\mathfrak{p} \in \mathrm{Spec}\, R \mid R/\mathfrak{p} \in \mathcal{X}\}.$$

(2) *If* $\varnothing \neq S \subseteq \mathrm{Spec}\, R$ *is specialization-closed*, *then*

$$\mathrm{Supp}^{-1}_{\mathsf{mod}\, R}\, S = \mathsf{thick}_{\mathsf{mod}\, R}\{R/\mathfrak{p} \mid \mathfrak{p} \in S\}.$$

Proof. (1) Let $\mathfrak{p}$ be a prime ideal. If X is a module in $\mathcal{X}$ whose support contains $\mathfrak{p}$, then $R/\mathfrak{p}$ is in the thick closure of k and X by Theorem 4.3, and hence $R/\mathfrak{p}$ is in $\mathcal{X}$. Conversely, if $R/\mathfrak{p}$ is in $\mathcal{X}$, then the support of $\mathcal{X}$ contains that of $R/\mathfrak{p}$, which contains $\mathfrak{p}$. Now the assertion follows.

(2) Let X be a module whose support is contained in S. Then the thick closure of $\{R/\mathfrak{p} \mid \mathfrak{p} \in S\}$ contains that of $\{k, R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp} X\}$, which contains X by Theorem 4.3. The set S contains $\mathrm{V}(\mathfrak{p}) = \operatorname{Supp}_R(R/\mathfrak{p})$ for each $\mathfrak{p} \in S$. Thus the assertion is shown. □

The following is the main result of this section, whose essential part is included in Theorem 4.3. Compare it with the similar results [Takahashi 2013a, Theorems 5.6 and 6.11] and [Takahashi 2013b, Theorem 5.1(2)].

Theorem 4.5. *Let $(R, \mathfrak{m}, k)$ be a local ring with $\dim R \le 2$ and having an isolated singularity.*

(1) *Every thick subcategory of $\operatorname{mod} R$ containing k is Serre.*

(2) *There is a one-to-one correspondence*

$$\left\{\begin{matrix}\textit{Thick subcategories of } \operatorname{mod} R\\ \textit{containing } k\end{matrix}\right\} \underset{g}{\overset{f}{\underset{\longleftarrow}{\overset{\longrightarrow}{\scriptstyle 1-1}}}} \left\{\begin{matrix}\textit{Specialization-closed subsets of } \operatorname{Spec} R\\ \textit{containing } \mathfrak{m}\end{matrix}\right\},$$

where f and g are defined by $f(\mathcal{X}) = \operatorname{Supp}_R \mathcal{X}$ and $g(S) = \operatorname{Supp}^{-1}_{\operatorname{mod} R} S$.

Proof. (1) Let $\mathcal{X}$ be a thick subcategory of $\operatorname{mod} R$ containing k. It suffices to show that $\mathcal{X} = \operatorname{Supp}^{-1}(\operatorname{Supp} \mathcal{X})$, because this equality especially says that $\mathcal{X}$ is a Serre subcategory. It is obvious that $\mathcal{X}$ is contained in $\operatorname{Supp}^{-1}(\operatorname{Supp} \mathcal{X})$. Let M be an R-module whose support is contained in that of $\mathcal{X}$. Take any prime ideal $\mathfrak{p}$ in the support of M. Then there exists an R-module $X \in \mathcal{X}$ whose support contains $\mathfrak{p}$. Theorem 4.3 implies that $R/\mathfrak{p}$ is in the thick closure of k and X, which is contained in $\mathcal{X}$. By Lemma 2.3(1) we see that M is in $\mathcal{X}$. Thus $\mathcal{X}$ contains $\operatorname{Supp}^{-1}(\operatorname{Supp} \mathcal{X})$, and the above equality follows.

(2) The assertion follows from (1) and Gabriel's classification [1962] of Serre subcategories. □

The assertion of Theorem 4.5 is no longer true for thick subcategories that do not contain k:

Example 4.6. Let R be a nonregular local ring with residue field k. Let $\mathcal{X}$ be the subcategory of $\operatorname{mod} R$ consisting of modules of finite projective dimension. Then $\mathcal{X}$ is a thick subcategory which does not contain k. There is an exact sequence $R \to k \to 0$, and we have $R \in \mathcal{X}$ and $k \notin \mathcal{X}$. This means that $\mathcal{X}$ is not a Serre subcategory of $\operatorname{mod} R$.

5. Rings of prime characteristic and Cohen–Macaulay modules

In this section, as in the prior two sections, we study the structure of $\operatorname{thick}_{\operatorname{mod} R}\{k, M\}$; we restrict ourselves to the case where R has prime characteristic and M is Cohen–Macaulay.

Let R be a ring of prime characteristic p, and let $q = p^e$ be a power of p. For a sequence $\boldsymbol{x} = x_1, \dots, x_n$ of elements of R we set $\boldsymbol{x}^q = x_1^q, \dots, x_n^q$. For an ideal I of R we denote by $I^{[q]}$ the ideal of R generated by the elements of the form a^q with $a \in I$. Note that if I is generated by a sequence $\boldsymbol{x}$ of elements of R, then $I^{[q]}$ is generated by the sequence $\boldsymbol{x}^q$. Note also that for each multiplicatively closed subset S of R one has $(IR_S)^{[q]} = I^{[q]}R_S$.

Lemma 5.1. *Let $(R, \mathfrak{m}, k)$ be a regular local ring of characteristic $p > 0$. Let $\boldsymbol{x} = x_1, \dots, x_d$ be a regular system of parameters of R. Let $q = p^e$ be a power of p. Let M be a nonzero R-module such that $\boldsymbol{x}^q M = 0$. Then there exists a nonzero free $R/(\boldsymbol{x}^q)$-module N possessing a filtration $0 = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_t = N$ in $\operatorname{mod} R$ with $N_i/N_{i-1} \cong M$ for $1 \le i \le t$.*

Proof. We regard M as an $R/(\boldsymbol{x}^q)$-module. Cohen's structure theorem implies that the completion $\widehat{R}$ of R is isomorphic to $k[[x_1, \dots, x_d]]$. As $R/(\boldsymbol{x}^q)$ is artinian, it is complete. There are isomorphisms of k-algebras

$$\begin{aligned} R/(\boldsymbol{x}^q) &\cong \widehat{R/(\boldsymbol{x}^q)} \cong \widehat{R}/\boldsymbol{x}^q\widehat{R} \\ &\cong k[[x_1, \dots, x_d]]/(x_1^q, \dots, x_d^q) = k[x_1, \dots, x_d]/(x_1^q, \dots, x_d^q) \cong kG, \end{aligned}$$

where kG denotes the group algebra of the finite abelian p-group $G = (\mathbb{Z}/q\mathbb{Z})^{\oplus d}$; see [Iyengar 2004, (1.4)]. Hence one can identify $R/(\boldsymbol{x}^q)$ with kG. The tensor product $N := M \otimes_k kG$ is a kG-module via the diagonal action, and is projective; see Theorem (3.2) of the same work. Since kG is a (commutative) local ring, N is a nonzero finitely generated free kG-module. Tensoring over k the composition series of kG with M, we have a filtration of N as in the assertion. $\square$

Denote by $\operatorname{fl} R$ the subcategory of $\operatorname{mod} R$ consisting of R-modules of finite length. Using the above lemma, we get a result on the structure of the thick closure of a finite length module.

Theorem 5.2. *Let R be a regular local ring of positive characteristic. Let M be a nonzero R-module of finite length. One then has $\operatorname{thick}_{\operatorname{mod} R} M = \operatorname{fl} R$.*

Proof. It is evident that the thick closure of M is contained in $\operatorname{fl} R$. As for the opposite inclusion relation, it is enough to show that the residue field k of R belongs to $\operatorname{thick}_{\operatorname{mod} R} M$. Let $\boldsymbol{x} = x_1, \dots, x_d$ be a regular system of parameters of R. Let p be the characteristic of R, and let $q = p^e$ be a power of p such that $\boldsymbol{x}^q M = 0$. Lemma 5.1 shows that there exists a nonzero free $R/(\boldsymbol{x}^q)$-module N having a filtration $0 = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_t = N$ in $\operatorname{mod} R$ with $N_i/N_{i-1} \cong M$ for $1 \le i \le t$. Note then that N is in the thick closure of M. We have only to show that k is in $\operatorname{thick}_{\operatorname{mod} R} R/(\boldsymbol{x}^q)$.

Let us do this by induction on $d = \dim R$. When $d = 0$, we have $R = k = R/(\boldsymbol{x}^q)$, and the statement trivially holds. Let $d > 0$, and put $\bar{R} = R/(x_1^q, \dots, x_{d-1}^q)$. There are exact sequences

$$\begin{aligned}
&0 \to \bar{R}/x_d^q\bar{R} \to \bar{R}/x_d^{q-1}\bar{R} \oplus \bar{R}/x_d^{q+1}\bar{R} \to \bar{R}/x_d^q\bar{R} \to 0,\\
&0 \to \bar{R}/x_d^{q-1}\bar{R} \to \bar{R}/x_d^{q-2}\bar{R} \oplus \bar{R}/x_d^{q}\bar{R} \to \bar{R}/x_d^{q-1}\bar{R} \to 0,\\
&\quad\vdots\\
&0 \to \bar{R}/x_d^2\bar{R} \to \bar{R}/x_d\bar{R} \oplus \bar{R}/x_d^3\bar{R} \to \bar{R}/x_d^2\bar{R} \to 0.
\end{aligned}$$

by Lemma 4.2(1). It can be seen from these exact sequences that $\bar{R}/x_d\bar{R}$ is in $\mathsf{thick}_{\mathsf{mod}\,R}\,\bar{R}/x_d^q\bar{R}$. Note $\bar{R}/x_d^q\bar{R} = R/(\boldsymbol{x}^q)$ and $\bar{R}/x_d\bar{R} = \widetilde{R}/(x_1^q,\dots,x_{d-1}^q)\widetilde{R}$, where $\widetilde{R} := R/(x_d)$. The induction hypothesis implies $k \in \mathsf{thick}_{\mathsf{mod}\,\widetilde{R}}\,\widetilde{R}/(x_1^q,\dots,x_{d-1}^q)\widetilde{R}$, and hence k is in $\mathsf{thick}_{\mathsf{mod}\,R}\,\widetilde{R}/(x_1^q,\dots,x_{d-1}^q)\widetilde{R}$. Consequently, we obtain $k \in \mathsf{thick}_{\mathsf{mod}\,R}\,R/(\boldsymbol{x}^q)$. □

Question 5.3. Does the assertion of Theorem 5.2 hold for any regular local ring R?

Remark 5.4. Using the Hopkins–Neeman theorem, one deduces that the derived category version of Theorem 5.2 holds: Let R be a regular ring. (We do not need to assume R is local or has prime characteristic.) Let $\mathsf{D}_{\mathsf{fl}}(R)$ stand for the subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ consisting of complexes having finite length homology. Let M be a nonzero object of $\mathsf{D}_{\mathsf{fl}}(R)$. Then $\mathsf{D}^{\mathsf{b}}(R) = \mathsf{D}_{\mathsf{perf}}(R)$ and $\operatorname{Supp} M = \operatorname{Supp} \mathsf{D}_{\mathsf{fl}}(R) = \{\mathfrak{m}\}$, whence by [Neeman 1992, Theorem 1.5] we have $\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\,M = \mathsf{D}_{\mathsf{fl}}(R)$.

Let R be a local ring with residue field k. Recall that an R-module M is called *Cohen–Macaulay* if $\operatorname{Ext}^i_R(k, M) = 0$ for all $i < \dim M$ (i.e., depth $M = \dim M$ or $M = 0$). Taking advantage of Lemma 5.1, we have the following similar theorem to Theorems 3.3 and 4.3:

Theorem 5.5. *Let $(R, \mathfrak{m}, k)$ be a local ring of prime characteristic p with an isolated singularity. Let $M \neq 0$ be a Cohen–Macaulay R-module. Then one has the equality*

$$\mathsf{thick}_{\mathsf{mod}\,R}\{k, M\} = \mathsf{thick}_{\mathsf{mod}\,R}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R M\}.$$

Proof. Lemma 2.3(1) and the fact $\mathfrak{m} \in \operatorname{Supp} M$ guarantee that the right-hand side contains the left-hand side. Let us show the opposite inclusion relation by induction on $\dim M$. When $\dim M = 0$, the module M has finite length, and we are done. Let $\dim M \geq 1$. We will be done if we prove that R/I is in $\mathcal{X} := \mathsf{thick}_{\mathsf{mod}\,R}\{k, M\}$ for all ideals I with $\mathrm{V}(I) \subseteq \operatorname{Supp} M$. Suppose that this does not hold, and let P be a maximal element (with respect to the inclusion relation) among the ideals I with $\mathrm{V}(I) \subseteq \operatorname{Supp} M$ and $R/I \notin \mathcal{X}$. Similarly to the proof of Theorem 3.3, the ideal P is a nonmaximal prime ideal belonging to the support of M, the module R/P is not in $\mathcal{X}$, and every R-module whose support is contained in $\mathrm{V}(P) - \{P\}$ belongs to $\mathcal{X}$. Since M is Cohen–Macaulay, we have $\operatorname{Ass} M = \operatorname{Min} M = \operatorname{Assh} M$ by [Bruns and Herzog 1998, Theorem 2.1.2(a)]. Suppose that P is not an associated prime of M. Then we find an M-regular element x in P. The exact sequence $0 \to M \xrightarrow{x} M \to M/xM \to 0$ shows that $\mathcal{X}$ contains $\mathsf{thick}_{\mathsf{mod}\,R}\{k, M/xM\}$. Note that M/xM is also a Cohen–Macaulay R-module whose support contains P. The induction hypothesis implies that R/P is in $\mathsf{thick}_{\mathsf{mod}\,R}\{k, M/xM\}$, and hence $\mathcal{X}$ contains R/P, which is a contradiction. Therefore P is an associated prime of M.

Let N be a P-primary component of the zero submodule 0 of M. Then $0 = N \cap L$ for some submodule L of M, which induces an exact sequence of R-modules $0 \to M \to M/N \oplus M/L \to M/N + L \to 0$. Observe that each prime ideal in the support of $M/N + L$ strictly contains P. Hence $\mathcal{X}$ contains $M/N + L$, and therefore $\mathcal{X}$ also contains M/N.

Since R_P is a regular local ring, by Lemma 3.1(1) one can choose a sequence $\boldsymbol{x} = x_1, \dots, x_n$ of elements in P with $\operatorname{ht} P = n = \operatorname{ht}(\boldsymbol{x})$ and $PR_P = \boldsymbol{x}R_P$. Note the equality $\operatorname{Ass} M/N = \{P\}$ implies the R_P-module $(M/N)_P$ has finite length. Applying Lemma 5.1, we see that for large enough $q = p^e$ there is a free $R_P/P^{[q]}R_P$-module Z of rank $r > 0$ possessing a filtration $0 = Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_t = Z$ in $\mathsf{mod}\,R_P$ with $Z_i/Z_{i-1} \cong (M/N)_P$ for $1 \le i \le t$. Using Lemma 4.2(2), one can inductively choose R-modules $W_0, \dots, W_t$ such that there exists a filtration $0 = W_0 \subsetneq W_1 \subsetneq \cdots \subsetneq W_t = W$ in $\mathsf{mod}\,R$ with $W_i/W_{i-1} \cong M/N$ for $1 \le i \le t$ and $W_P \cong Z \cong (R_P/P^{[q]}R_P)^{\oplus r}$. Then W is in the thick closure of M/N, and hence in $\mathcal{X}$. There is an exact sequence $0 \to K \to W \to (R/P^{[q]})^{\oplus r} \to C \to 0$ with $K_P = 0 = C_P$. Note that $\operatorname{Supp} W = \operatorname{Supp} M/N = \mathrm{V}(P) = \operatorname{Supp}(R/P^{[q]})^{\oplus r}$. Hence the supports of K, C are contained in $\mathrm{V}(P) - \{P\}$, which implies $\mathcal{X}$ contains K and C. Therefore the module $R/P^{[q]}$ is in $\mathcal{X}$.

There is an exact sequence

$$0 \to R_P/\boldsymbol{x}^q R_P \to (R_P/\mathfrak{a}_1 R_P) \oplus (R_P/\mathfrak{a}_2 R_P) \to R_P/\boldsymbol{x}^q R_P \to 0$$

by Lemma 4.2(1), where $\mathfrak{a}_1 = (x_1^q, \dots, x_{n-1}^q, x_n^{q-1})R$ and $\mathfrak{a}_2 = (x_1^q, \dots, x_{n-1}^q, x_n^{q+1})R$. Put $\mathfrak{b}_i = \mathfrak{a}_i R_P \cap R$ for $i = 1, 2$. Since P is a minimal prime of $\mathfrak{a}_i$, the ideal $\mathfrak{b}_i$ is the P-primary component of $\mathfrak{a}_i$. Note that $\mathfrak{a}_i R_P = \mathfrak{b}_i R_P$, and $\mathrm{V}(\mathfrak{b}_i) = \mathrm{V}(\sqrt{\mathfrak{b}_i}) = \mathrm{V}(P)$ for $i = 1, 2$. Setting $E = R/\mathfrak{b}_1 \oplus R/\mathfrak{b}_2$, we see from Lemma 4.2(2) that there is an exact sequence $0 \to R/P^{[q]} \to U \to R/P^{[q]} \to 0$ such that $U_P \cong E_P$. We have $\operatorname{Supp} E = \mathrm{V}(\mathfrak{b}_1) \cup \mathrm{V}(\mathfrak{b}_2) = \mathrm{V}(P) = \operatorname{Supp} U$. Choosing an exact sequence $0 \to K' \to U \to E \to C' \to 0$ with $K'_P = 0 = C'_P$, we see that the supports of K' and C' are contained in $\mathrm{V}(P) - \{P\}$, whence they are in $\mathcal{X}$. As U is in $\mathcal{X}$, so is E, and so is $R/\mathfrak{b}_1$.

Since $(R/\mathfrak{b}_1)_P = R_P/(x_1^q, \dots, x_{n-1}^q, x_n^{q-1})R_P$, the same argument as above shows that $R/\mathfrak{c}$ belongs to $\mathcal{X}$ with $\mathfrak{c} = (x_1^q, \dots, x_{n-1}^q, x_n^{q-2})R_P \cap R$ if $q > 2$. Iterating this procedure yields that $R/(x_1, \dots, x_n)R_P \cap R$ belongs to $\mathcal{X}$. (Here we use the fact that any permutation of a regular sequence on a local ring is again regular.) Since $(x_1, \dots, x_n)R_P \cap R = PR_P \cap R = P$, this means that R/P is in $\mathcal{X}$, which is a contradiction. This completes the proof of the theorem. $\square$

6. Thick subcategories of derived categories containing the residue field

From this section to the end of this paper, we deal with thick subcategories of derived categories. In this section, we prove a classification theorem of thick subcategories containing the residue field over an isolated singularity.

We begin with a well-known statement. In view of this, it is reasonable to think of classifying, for a general local ring R, the thick subcategories of $\mathsf{D^b}(R)$ containing the residue field.

Remark 6.1. Let R be a local ring with residue field k. The following are equivalent.

(1) The ring R is regular.
(2) Every nonzero thick subcategory of $\mathsf{D^b}(R)$ contains k.
(3) For each nonzero object X of $\mathsf{D^b}(R)$, the thick closure of X contains k.

The following lemma helps us make the derived category version of Theorem 3.3:

Lemma 6.2. *Let R be an isolated singularity. Let X be a bounded complex of R-modules. Then X is quasi-isomorphic to a complex*

$$Y = (0 \to Y^s \to Y^{s+1} \to \cdots \to Y^t \to 0)$$

with $s \le t$ such that Y^i is free for all $s+1 \le i \le t$ and Y^s is locally free on the punctured spectrum of R.

Proof. Take a free resolution $F = (\cdots \xrightarrow{\delta^{t-2}} F^{t-1} \xrightarrow{\delta^{t-1}} F^t \to 0)$ of X. Choose an integer u such that $\mathrm{H}^i(F) = 0$ for all $i < u$, and put $d = \dim R$. Then $C := \operatorname{Cok} \delta^{u-d-1}$ is a d-th syzygy of $\operatorname{Cok} \delta^{u-1}$, which is locally free on the punctured spectrum. The complex X is quasi-isomorphic to the complex $(0 \to C \to F^{u-d+1} \to \cdots \to F^t \to 0)$. $\square$

Now we can prove the following theorem analogous to Theorems 3.3, 4.3 and 5.5.

Theorem 6.3. *Let $(R, \mathfrak{m}, k)$ be a local ring with an isolated singularity. Let X be a nonacyclic bounded complex of R-modules. Then one has*

$$\mathsf{thick}_{\mathsf{D^b}(R)}\{k, X\} = \mathsf{thick}_{\mathsf{D^b}(R)}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R X\}.$$

Proof. The inclusion $(\subseteq)$ follows from Lemma 2.3(1)(3) and the fact $\mathfrak{m} \in \operatorname{Supp} X$. Let us show the opposite inclusion $(\supseteq)$. By Lemma 6.2 we may assume that X has the form $X = (0 \to X^s \to X^{s+1} \to \cdots \to X^t \to 0)$ such that the R-module X^i is free for all $s+1 \le i \le t$ and X^s is locally free on the punctured spectrum of R. Set $\mathcal{X} = \mathsf{thick}_{\mathsf{D^b}(R)}\{k, X\}$. It suffices to prove $R/I \in \mathcal{X}$ for all ideals I of R with $\mathrm{V}(I) \subseteq \operatorname{Supp} X$. Similarly to the proof of Theorem 3.3, we show this by contradiction. Assume that this does not hold, and choose a maximal element P of the set of ideals I with $R/I \notin \mathcal{X}$ and $\mathrm{V}(I) \subseteq \operatorname{Supp} X$. We then have:

Claim. (1) One has $\mathfrak{m} \neq P \in \operatorname{Supp} X$ and $R/P \notin \mathcal{X}$.

(2) Let C be an object of $\mathsf{D^b}(R)$. If $\operatorname{Supp}_R C$ is contained in $\mathrm{V}(P) - \{P\}$, then C is in $\mathcal{X}$.

Proof of Claim. (1) This is similarly shown to the Claim in the proof of Theorem 3.3.

(2) Take any $\mathfrak{p} \in \operatorname{Supp}_R \mathrm{H}(C) = \operatorname{Supp}_R C$. Then $\mathfrak{p}$ strictly contains P, and the maximality of P implies $R/\mathfrak{p} \in \mathcal{X}$. By Lemma 2.3(1), $\mathrm{H}(C)$ is in $\mathcal{X}$. Lemma 2.3(3) shows $C \in \mathcal{X}$. $\square$

Since R is an isolated singularity, R_P is regular. By Lemma 3.1, there exists an exact sequence

$$0 \to R/(\boldsymbol{x}) \to R/P \oplus R/Q \to R/I \to 0, \tag{6.3.1}$$

where $\boldsymbol{x} = x_1, \dots, x_n$ is a sequence in R generating an ideal of height n, and I is an ideal strictly containing P. Let

$$K = \mathrm{K}(\boldsymbol{x}, X) = (0 \to X^{\oplus\binom{n}{n}} \to X^{\oplus\binom{n}{n-1}} \to \cdots \to X^{\oplus\binom{n}{1}} \to X^{\oplus\binom{n}{0}} \to 0)$$

be the Koszul complex of $\boldsymbol{x}$ on X, which is a complex of objects of the abelian category $\mathsf{C^b}(R)$. Put $\mathcal{Y} = \mathsf{thick}_{\mathsf{C^b}(R)}\{k, X\}$. For each integer i the i-th homology $\mathrm{H}_i(K)$ of K is the complex $(0 \to \mathrm{H}_i(\boldsymbol{x}, X^s) \to \mathrm{H}_i(\boldsymbol{x}, X^{s+1}) \to \cdots \to \mathrm{H}_i(\boldsymbol{x}, X^t) \to 0)$ of R-modules, where $\mathrm{H}_i(\boldsymbol{x}, X^j)$ stands for the (usual) i-th Koszul homology of $\boldsymbol{x}$ on the R-module X^j. Lemma 3.2 implies that $\mathrm{H}_i(\boldsymbol{x}, X^j)$ has finite length for each $i > 0$ and $s \le j \le t$. By Lemma 2.3(4) we observe that $\mathrm{H}_i(K)$ belongs to $\mathcal{Y}$ for every $i > 0$, and by Lemma 2.3(5) the complex $\mathrm{H}_0(K) = R/(\boldsymbol{x}) \otimes_R X$ also belongs to $\mathcal{Y}$. Consequently, the complex $R/(\boldsymbol{x}) \otimes_R X$, as an object of $\mathsf{D^b}(R)$, is in $\mathcal{X}$.

The short exact sequence (6.3.1) induces an exact sequence

$$\operatorname{Tor}_1^R(R/I, X) \xrightarrow{f} R/(\boldsymbol{x}) \otimes_R X \to (R/P \otimes_R X) \oplus (R/Q \otimes_R X) \to R/I \otimes_R X \to 0$$

in the abelian category $\mathsf{C^b}(R)$, where $\operatorname{Tor}_1^R(R/I, X)$ stands for the induced complex $(0 \to \operatorname{Tor}_1^R(R/I, X^s) \to \operatorname{Tor}_1^R(R/I, X^{s+1}) \to \cdots \to \operatorname{Tor}_1^R(R/I, X^t) \to 0)$ of R-modules. Let Z be the image of the morphism f. Note that each component Z^i of the complex Z is a homomorphic image of the R-module $\operatorname{Tor}_1^R(R/I, X^i)$. Hence one has $\operatorname{Supp} Z^i \subseteq \mathrm{V}(I) \subseteq \mathrm{V}(P) - \{P\}$ for each i, and therefore $\operatorname{Supp} Z = \bigcup_{i \in \mathbb{Z}} \operatorname{Supp} \mathrm{H}^i(Z) \subseteq \bigcup_{i \in \mathbb{Z}} \operatorname{Supp} Z^i \subseteq \mathrm{V}(P) - \{P\}$. It follows from the Claim that Z, as an object of $\mathsf{D^b}(R)$, belongs to $\mathcal{X}$. Similarly, it is seen that $R/I \otimes_R X \in \mathcal{X}$. The induced exact sequence

$$0 \to Z \to R/(\boldsymbol{x}) \otimes_R X \to (R/P \otimes_R X) \oplus (R/Q \otimes_R X) \to R/I \otimes_R X \to 0$$

shows that the subcategory $\mathcal{X}$ of $\mathsf{D}^{\mathsf{b}}(R)$ contains the complex

$$R/P \otimes_R X = (0 \to X^s/PX^s \to X^{s+1}/PX^{s+1} \to \cdots \to X^t/PX^t \to 0).$$

Since X^s/PX^s is a finitely generated module over the integral domain R/P, it has a rank, say r. There is an exact sequence $0 \to (R/P)^{\oplus r} \to X^s/PX^s \to C \to 0$ of R/P-modules such that $\dim C < \dim R/P$. We obtain a short exact sequence $0 \to W \to R/P \otimes_R X \to C[-s] \to 0$ in $\mathsf{C}^{\mathsf{b}}(R)$, where

$$W = (0 \to (R/P)^{\oplus r} \to X^{s+1}/PX^{s+1} \to \cdots \to X^t/PX^t \to 0).$$

As $PC = 0 = C_P$, the set $\operatorname{Supp}_R(C[-s]) = \operatorname{Supp}_R C$ is contained in $\mathrm{V}(P) - \{P\}$. The Claim yields that $C[-s]$ is in $\mathcal{X}$, and the above short exact sequence shows that W is in $\mathcal{X}$.

Note that W is a perfect complex of R/P-modules, and hence as an object of $\mathsf{D}^{\mathsf{b}}(R/P)$ it belongs to $\mathsf{D}_{\mathsf{perf}}(R/P)$. Since $C[-s]_P = 0$, we have isomorphisms $W_P = (R/P \otimes_R X)_P \cong \kappa(P) \otimes_{R_P} X_P \cong \kappa(P) \otimes^{\boldsymbol{L}}_{R_P} X_P$ in $\mathsf{D}^{\mathsf{b}}(R_P)$, where the last isomorphism follows from the fact that X_P is a perfect complex of R_P-modules. As P is in $\operatorname{Supp}_R X$, the complex W_P is not acyclic. This means that $\operatorname{Supp}_{R/P} W$ contains the zero ideal of R/P, and we obtain $\operatorname{Supp}_{R/P} W = \operatorname{Spec} R/P = \operatorname{Supp}_{R/P}(R/P)$. By [Neeman 1992, Theorem 1.5], R/P is in $\mathsf{thick}_{\mathsf{D}_{\mathsf{perf}}(R/P)} W = \mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R/P)} W$, and therefore it belongs to $\mathcal{X}$. This contradiction completes the proof. □

Remark 6.4. Similarly to Remark 3.4, the equality in Theorem 6.3 is no longer true if we remove k from the left-hand side; the equality

$$\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)} X = \mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\{R/\mathfrak{p} \mid \mathfrak{p} \in \operatorname{Supp}_R X\}$$

holds for $X = R$ if and only if $\mathsf{D}_{\mathsf{perf}}(R) = \mathsf{D}^{\mathsf{b}}(R)$, if and only if R is regular. This is one of the reasons why we consider thick subcategories containing k.

Using Theorem 6.3, we get a derived category version of Corollary 4.4:

Corollary 6.5. *Let $(R, \mathfrak{m}, k)$ be a local ring with an isolated singularity.*

(1) *If $\mathcal{X}$ is a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ containing k, then*

$$\operatorname{Supp}_R \mathcal{X} = \{\mathfrak{p} \in \operatorname{Spec} R \mid R/\mathfrak{p} \in \mathcal{X}\}.$$

(2) *If $S \neq \varnothing$ is a specialization-closed subset of $\operatorname{Spec} R$, then*

$$\operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)} S = \mathsf{thick}\{R/\mathfrak{p} \mid \mathfrak{p} \in S\}.$$

The following is the main theorem of this section:

Theorem 6.6. *Let $(R, \mathfrak{m}, k)$ be a local ring with an isolated singularity. The assignments $f : \mathcal{X} \mapsto \operatorname{Supp}_R \mathcal{X}$ and $g : S \mapsto \operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)} S$ make mutually inverse bijections*

$$\left\{\begin{matrix}\textit{Thick subcategories of } \mathsf{D}^{\mathsf{b}}(R) \\ \textit{containing } k\end{matrix}\right\} \underset{g}{\overset{f}{\underset{\leftarrow}{\overset{\rightarrow}{\scriptstyle 1-1}}}} \left\{\begin{matrix}\textit{Specialization-closed subsets of } \operatorname{Spec} R \\ \textit{containing } \mathfrak{m}\end{matrix}\right\}.$$

Proof. Let $\mathcal{X}$ be a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ containing k, and let S be a specialization-closed subset of $\operatorname{Spec} R$ containing $\mathfrak{m}$.

(1) The set $\operatorname{Supp} \mathcal{X}$ is specialization-closed. Since the residue field k is in $\mathcal{X}$, the maximal ideal $\mathfrak{m}$ is in the support of $\mathcal{X}$. Hence f is a well-defined map.

(2) The subcategory $\operatorname{Supp}^{-1} S$ is thick. The support of k is contained in $\{\mathfrak{m}\}$, which is contained in S. Hence k is in $\operatorname{Supp}^{-1} S$, and g is a well-defined map.

(3) It is obvious that $\operatorname{Supp}\operatorname{Supp}^{-1} S$ is contained in S. Let $\mathfrak{p}$ be a prime ideal in S. We have $\operatorname{Supp}_R R/\mathfrak{p} = \mathrm{V}(\mathfrak{p})$, which is contained in S as S is specialization-closed. Hence $\mathfrak{p}$ is in $\operatorname{Supp}_R R/\mathfrak{p}$ and $R/\mathfrak{p}$ is in $\operatorname{Supp}^{-1} S$. Thus we obtain $S = \operatorname{Supp}\operatorname{Supp}^{-1} S$.

(4) Clearly, the subcategory $\operatorname{Supp}^{-1}\operatorname{Supp} \mathcal{X}$ contains $\mathcal{X}$. Let C be an object of $\mathsf{D}^{\mathsf{b}}(R)$ whose support is contained in that of $\mathcal{X}$. Take a prime ideal $\mathfrak{p} \in \operatorname{Supp} C$. Then $\mathfrak{p}$ is in the support of X for some $X \in \mathcal{X}$. Theorem 6.3 implies that $R/\mathfrak{p}$ belongs to the thick closure of k and X, which is contained in $\mathcal{X}$. Thus $R/\mathfrak{p}$ is in $\mathcal{X}$ for all prime ideals $\mathfrak{p}$ in the support of C. Using Theorem 6.3 again, we observe that C belongs to $\mathcal{X}$. Consequently, we obtain $\mathcal{X} = \operatorname{Supp}^{-1}\operatorname{Supp} \mathcal{X}$.

Getting the above **(1)**–**(4)** together completes the proof of the theorem. □

Remark 6.7. An anonymous referee has pointed out that Theorem 6.6 can also be shown as follows: Let $U = \operatorname{Spec} R \setminus \{\mathfrak{m}\}$ be the punctured spectrum of R. The assumption that R has an isolated singularity implies that U is a regular scheme. On one hand, by [Thomason 1997, Theorem 3.15] the thick subcategories of $\mathsf{D}^{\mathsf{b}}(\operatorname{coh} U)$ correspond to the specialization-closed subsets of U, which are the same as the specialization-closed subsets of $\operatorname{Spec} R$ containing $\mathfrak{m}$. On the other hand, since $\mathsf{D}^{\mathsf{b}}(\operatorname{coh} U)$ is equivalent to $\mathsf{D}^{\mathsf{b}}(R)/\operatorname{thick} k$ by [Orlov 2011, Lemma 2.2], the thick subcategories of $\mathsf{D}^{\mathsf{b}}(\operatorname{coh} U)$ correspond to the thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$ containing k.

This is a simpler proof, using techniques in algebraic geometry. Our methods are purely ring-theoretic, and also essentially the same as those in the proof of Theorem 3.3, for which the approach the referee mentions does not seem to work. It is thus worth giving our methods.

Unless R has an isolated singularity, Theorem 6.6 does not necessarily hold:

Remark 6.8. Let $(R, \mathfrak{m}, k)$ be a local ring, and suppose that R does not have an isolated singularity. Set $\mathcal{X} = \mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\{k, R\}$. Then $\mathcal{X}$ is a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ containing k, but $\mathcal{X} \neq \operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)} S$ for all subsets S of $\operatorname{Spec} R$.

One has a classification theorem of thick subcategories without using prime ideals:

Corollary 6.9. *If R is a local ring with an isolated singularity, one has a* 1-1 *correspondence*

$$\left\{\begin{matrix}\textit{Thick subcategories of } \mathsf{D}^{\mathsf{b}}(R)\\ \textit{containing } k\end{matrix}\right\} \underset{\psi}{\overset{\phi}{\underset{1-1}{\rightleftarrows}}} \left\{\begin{matrix}\textit{Nonzero thick subcategories}\\ \textit{of } \mathsf{D}_{\mathsf{perf}}(R)\end{matrix}\right\},$$

where ϕ, ψ are defined by $\phi(\mathcal{X}) = \mathcal{X} \cap \mathsf{D}_{\mathsf{perf}}(R)$ and $\psi(\mathcal{Y}) = \mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}(\mathcal{Y} \cup \{k\})$ for subcategories $\mathcal{X}$ of $\mathsf{D}^{\mathsf{b}}(R)$ and $\mathcal{Y}$ of $\mathsf{D}_{\mathsf{perf}}(R)$.

Proof. Let S be a specialization-closed subset of $\operatorname{Spec} R$ containing $\mathfrak{m}$. Take a system of generators $\boldsymbol{x}$ of $\mathfrak{m}$. Then $\operatorname{Supp}^{-1}_{\mathsf{D}_{\mathsf{perf}}(R)} S$ contains the Koszul complex $\mathrm{K}(\boldsymbol{x}, R)$, and hence it is a nonzero thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$. Conversely, for any nonzero thick subcategory $\mathcal{Y}$ of $\mathsf{D}_{\mathsf{perf}}(R)$, the support $\operatorname{Supp}_R \mathcal{Y}$ contains $\mathfrak{m}$. Thus, [Neeman 1992, Theorem 1.5] implies that Supp_R and $\operatorname{Supp}^{-1}_{\mathsf{D}_{\mathsf{perf}}(R)}$ make mutually inverse bijections between the nonzero thick subcategories of $\mathsf{D}_{\mathsf{perf}}(R)$ and the specialization-closed subsets of $\operatorname{Spec} R$ containing $\mathfrak{m}$.

Let $\mathcal{X}$ be a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ containing k, and let $\mathcal{Y}$ be a nonzero thick subcategory of $\mathsf{D}_{\mathsf{perf}}(R)$. Combining our Theorem 6.6 with the above one-to-one correspondence, one has only to verify the equalities

(1) $\operatorname{Supp}^{-1}_{\mathsf{D}_{\mathsf{perf}}(R)} \operatorname{Supp} \mathcal{X} = \mathcal{X} \cap \mathsf{D}_{\mathsf{perf}}(R)$,

(2) $\operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)} \operatorname{Supp} \mathcal{Y} = \mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}(\mathcal{Y} \cup \{k\})$.

We have $\mathcal{X} \cap \mathsf{D}_{\mathsf{perf}}(R) \subseteq \operatorname{Supp}^{-1}_{\mathsf{D}_{\mathsf{perf}}(R)}(\operatorname{Supp} \mathcal{X}) \subseteq \operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)}(\operatorname{Supp} \mathcal{X}) = \mathcal{X}$, where the last equality follows from Theorem 6.6. This shows (1). On the other hand, it holds that $\operatorname{Supp} \mathcal{Y} = \operatorname{Supp}(\mathcal{Y} \cup \{k\}) = \operatorname{Supp}(\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}(\mathcal{Y} \cup \{k\}))$, where the second equality follows from the fact that $\mathcal{Y}$ is nonzero. Applying $\operatorname{Supp}^{-1}_{\mathsf{D}^{\mathsf{b}}(R)}$ and using Theorem 6.6, we obtain (2). ☐

7. Hypersurfaces and Cohen–Macaulay rings with minimal multiplicity

In this section, using the classification obtained in the previous section, we explore thick subcategories over hypersurfaces and Cohen–Macaulay rings with minimal multiplicity.

Definition 7.1. (1) A local ring R is called a *hypersurface* if the completion of R is isomorphic to a quotient of a regular local ring by a nonzero element.

(2) Let R be a Cohen–Macaulay local ring. Then R satisfies the inequality

$$\mathrm{e}(R) \geq \operatorname{edim} R - \dim R + 1, \tag{7.1.1}$$

where $\mathrm{e}(R)$ and $\operatorname{edim} R$ denote the multiplicity of R and the embedding dimension of R, respectively. We say that R has *minimal multiplicity* (or *maximal embedding dimension*) if the equality of (7.1.1) holds.

(3) Let A_1, A_2 be sets whose intersection is possibly nonempty. The *disjoint union* of A_1 and A_2 is defined as

$$A_1 \sqcup A_2 = (A_1 \times \{1\}) \cup (A_2 \times \{2\}) = \{(x, 1), (y, 2) \mid x \in A_1, y \in A_2\}.$$

In the case where $A_1 \cap A_2$ is empty, the set $A_1 \sqcup A_2$ is identified with the union $A_1 \cup A_2$, namely, it is the usual disjoint union.

Below is the main result of this section. See Section 1 for the definition of standardness.

Theorem 7.2. *Let R be a nonregular local ring with an isolated singularity, which is either*

(1) *a hypersurface, or*

(2) *a Cohen–Macaulay ring with minimal multiplicity and infinite residue field.*

Then there is a one-to-one correspondence

$$\left\{\begin{matrix}\text{Standard thick}\\ \text{subcategories}\\ \text{of } \mathsf{D}^{\mathrm{b}}(R)\end{matrix}\right\} \underset{\Gamma}{\overset{\Lambda}{\underset{1-1}{\rightleftarrows}}} \left\{\begin{matrix}\text{Nonempty}\\ \text{specialization-closed}\\ \text{subsets of } \operatorname{Spec} R\end{matrix}\right\} \sqcup \left\{\begin{matrix}\text{Nonempty}\\ \text{specialization-closed}\\ \text{subsets of } \operatorname{Spec} R\end{matrix}\right\}.$$

Here, the maps Λ and Γ are defined by:

$$\Lambda(\mathcal{X}) = \begin{cases} (\operatorname{Supp} \mathcal{X}, 1) & \text{if } \mathcal{X} \subseteq \mathsf{D}_{\mathsf{perf}}(R), \\ (\operatorname{Supp} \mathcal{X}, 2) & \text{if } \mathcal{X} \nsubseteq \mathsf{D}_{\mathsf{perf}}(R), \end{cases}$$
$$\Gamma((S, i)) = \begin{cases} (\operatorname{Supp}^{-1} S) \cap \mathsf{D}_{\mathsf{perf}}(R) & \text{if } i = 1, \\ \operatorname{Supp}^{-1} S & \text{if } i = 2. \end{cases}$$

We shall give a proof of this theorem at the end of this section, after preparing several necessary tools. Here are some examples of a ring satisfying the assumption of Theorem 7.2(2).

Example 7.3. Let k be an infinite field, and let x, y be indeterminates over k. Then it is easy to observe that $k[[x, y]]/(x^2, xy, y^2)$, $k[[x, y, z]]/(x^2-yz, y^2-zx, z^2-xy)$ and $k[[x^3, x^2y, xy^2, y^3]]$ are non-Gorenstein rings satisfying the condition (2) in Theorem 7.2. In general, normal local rings of dimension two with rational singularities satisfy Theorem 7.2(2); see [Huneke and Watanabe 2015, Theorem 3.1].

Remark 7.4. (1) Theorem 7.2(1) can also be deduced from [Stevenson 2014, Theorem 4.9].

(2) Theorem 7.2(2) especially says the following.

> *Let $(R, \mathfrak{m}, k)$ be a Cohen–Macaulay local ring with an isolated singularity, and assume k is infinite and R has minimal multiplicity. Let $\mathcal{X}$ be a standard thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ which is not contained in $\mathsf{D}_{\mathsf{perf}}(R)$. Then $\mathcal{X}$ contains k.*

This statement is no longer true without the assumption that R has minimal multiplicity. Indeed, let $R = k[x, y]/(x^2, y^2)$ with k a field, and let $\mathcal{X}$ be the thick closure of R and $R/(x)$ in $\mathsf{D}^{\mathsf{b}}(R)$. Then R is an artinian complete intersection local ring, and $\mathcal{X}$ is a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$. As $R \in \mathcal{X}$, it is standard. Since $R/(x)$ has infinite projective dimension as an R-module, $\mathcal{X}$ is not contained in $\mathsf{D}_{\mathsf{perf}}(R)$. Both R and $R/(x)$ have complexity at most 1, and the subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ consisting of objects having complexity at most 1 is thick. Hence any object in $\mathcal{X}$ has complexity at most 1. The fact that k has complexity 2 shows $k \notin \mathcal{X}$.

Thus, the assumption in Theorem 7.2(2) that R has minimal multiplicity is indispensable.

We state a general lemma on triangulated categories, whose proof is standard and omitted.

Lemma 7.5. *Let $\mathcal{T}$ be an essentially small triangulated category.*

(1) *Let $\mathcal{U}$ be a thick subcategory of $\mathcal{T}$. Let $\pi : \mathcal{T} \to \mathcal{T}/\mathcal{U}$ be the canonical functor. Let T be an object of $\mathcal{T}$ and $\mathcal{X}$ a subcategory of $\mathcal{T}$. Then T is in* $\mathsf{thick}_{\mathcal{T}}(\mathcal{U} \cup \mathcal{X})$ *if and only if πT is in* $\mathsf{thick}_{\mathcal{T}/\mathcal{U}}(\pi\mathcal{X})$.

(2) *Let $\mathcal{C}$ be a subcategory of $\mathcal{T}$. For each object $T \in \mathsf{thick}_{\mathcal{T}}\,\mathcal{C}$ there exist a finite number of objects $C_1, \dots, C_n \in \mathcal{C}$ such that $T \in \mathsf{thick}_{\mathcal{T}}\{C_1, \dots, C_n\}$.*

The *stable derived category* $\mathsf{D}_{\mathsf{sg}}(R)$ of R, which is also called the *singularity category* of R, is defined as the Verdier quotient of $\mathsf{D}^{\mathsf{b}}(R)$ by $\mathsf{D}_{\mathsf{perf}}(R)$. The following proposition says that a standard thick subcategory generating the singularity category contains the residue field.

Proposition 7.6. *Let R be a local ring with residue field k. Let $\mathcal{X}$ be a standard thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$. Suppose that the equality* $\mathsf{thick}_{\mathsf{D}_{\mathsf{sg}}(R)}(\pi\mathcal{X}) = \mathsf{D}_{\mathsf{sg}}(R)$ *holds, where $\pi : \mathsf{D}^{\mathsf{b}}(R) \to \mathsf{D}_{\mathsf{sg}}(R)$ stands for the canonical functor. Then $\mathcal{X}$ contains k.*

Proof. Lemma 7.5(1) implies $\mathsf{thick}_{\mathsf{D^b}(R)}(\{R\} \cup \mathcal{X}) = \mathsf{D^b}(R)$. By Lemma 7.5(2) there is an object $X \in \mathcal{X}$ such that k belongs to the thick closure of R and X. Since $\mathcal{X}$ is standard, it contains a nonacyclic perfect complex P. Tensoring P shows that $P \otimes_R^{\boldsymbol{L}} k$ belongs to the thick closure of P and $P \otimes_R^{\boldsymbol{L}} X$, which is contained in $\mathcal{X}$. As P is not acyclic, the maximal ideal is in the support of P in $\mathsf{D^b}(R)$, which means $P \otimes_R^{\boldsymbol{L}} k \neq 0$ in $\mathsf{D^b}(R)$. Thus $P \otimes_R^{\boldsymbol{L}} k$ contains $k[n]$ as a direct summand for some $n \in \mathbb{Z}$, and it follows that k is in $\mathcal{X}$. □

For every triangulated category $\mathcal{T}$, the zero subcategory $\mathbf{0}$ and the whole category $\mathcal{T}$ are thick subcategories of $\mathcal{T}$. We call these two thick subcategories *trivial*, and the other thick subcategories *nontrivial*. The assumption of Theorem 7.2 comes from the fact that the following proposition holds under it.

Proposition 7.7. *Let R be a local ring with an isolated singularity. Suppose that R is either*

(1) *a hypersurface*, *or*

(2) *a Cohen–Macaulay ring with minimal multiplicity and infinite residue field.*

Then $\mathsf{D_{sg}}(R)$ has no nontrivial thick subcategory.

Proof. (1) By virtue of [Takahashi 2010, Main Theorem], the thick subcategories of $\mathsf{D_{sg}}(R)$ bijectively correspond to the specialization-closed subsets of the singular locus $\operatorname{Sing} R$ of R, i.e., the set of prime ideals $\mathfrak{p}$ of R such that the local ring $R_{\mathfrak{p}}$ is not regular. Since R has an isolated singularity, $\operatorname{Sing} R$ is trivial. Thus there exist only trivial thick subcategories of $\mathsf{D_{sg}}(R)$.

(2) Let $\mathcal{X}$ be a nonzero thick subcategory of $\mathsf{D_{sg}}(R)$. Then there exists a bounded R-complex C of infinite projective dimension such that πC is in $\mathcal{X}$, where $\pi : \mathsf{D^b}(R) \to \mathsf{D_{sg}}(R)$ is the canonical functor. One finds an exact triangle $P \to C \to M[n] \rightsquigarrow$ in $\mathsf{D^b}(R)$ with $P \in \mathsf{D_{perf}}(R)$ and $n \in \mathbb{Z}$ such that M is the $(d+1)$-st syzygy of an R-module, where $d = \dim R$. As C has infinite projective dimension, M is a nonzero module. The object πC is isomorphic to $\pi M[n]$ in $\mathsf{D_{sg}}(R)$, whence πM belongs to $\mathcal{X}$.

There is a maximal Cohen–Macaulay R-module N such that $M \cong \Omega_R N$. Since R has minimal multiplicity and the residue field of R is infinite, we find a parameter ideal $Q = (x_1, \dots, x_d)$ of R such that $\mathfrak{m}^2 = Q\mathfrak{m}$; see [Bruns and Herzog 1998, Exercise 4.6.14]. Note that $\boldsymbol{x} := x_1, \dots, x_d$ is a regular sequence on R, and hence on N. We see that M/QM is isomorphic to $\Omega_{R/Q}(N/QN)$, which is contained in $\mathfrak{m}L$ for some free R/Q-module L. Since $\mathfrak{m}^2$ is contained in Q, the module $\Omega_{R/Q}(N/QN)$ is annihilated by $\mathfrak{m}$, which implies that M/QM is a nonzero k-vector space.

In the derived category $\mathsf{D^b}(R)$ the module M/QM is isomorphic to the Koszul complex $K := \mathrm{K}(\boldsymbol{x}, M)$. Since K is a bounded complex of direct sums of copies of M, the object πK belongs to the thick closure of πM (see Lemma 2.3(4)), and hence πK belongs to $\mathcal{X}$. Consequently, the object πk is in $\mathcal{X}$. As R has an isolated singularity, $\mathsf{D}_{\mathsf{sg}}(R)$ coincides with the thick closure of πk by Corollary 3.5. This implies $\mathcal{X} = \mathsf{D}_{\mathsf{sg}}(R)$, which is what we want. □

We give a lemma on elementary set theory, whose proof is also elementary and omitted.

Lemma 7.8. *Let A_1, A_2, B_1, B_2 be sets. Let $f_i : A_i \to B_i$ be a bijection for each $i = 1, 2$. Define the map $g : A_1 \sqcup A_2 \to B_1 \sqcup B_2$ by $g((a, i)) = (f_i(a), i)$ for $a \in A_i$ and $i = 1, 2$. Then g is a bijection.*

Now we can prove Theorem 7.2:

Proof of Theorem 7.2. Let A_1 be the set of nonzero thick subcategories of $\mathsf{D}_{\mathsf{perf}}(R)$. Let A_2 be the set of standard thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$ not contained in $\mathsf{D}_{\mathsf{perf}}(R)$. Then $A_1 \cap A_2$ is empty, and $A_1 \sqcup A_2$ coincides with the set of standard thick subcategories of $\mathsf{D}^{\mathsf{b}}(R)$. Let B be the set of nonempty specialization-closed subsets of Spec R. By [Neeman 1992, Theorem 1.5] there is a one-to-one correspondence $f : A_1 \rightleftarrows B : g$ defined by $f(\mathcal{X}) = \operatorname{Supp} \mathcal{X}$ and $g(S) = (\operatorname{Supp}^{-1} S) \cap \mathsf{D}_{\mathsf{perf}}(R)$. In view of Lemma 7.8, it suffices to show that there is a one-to-one correspondence $p : A_2 \rightleftarrows B : q$ defined by $p(\mathcal{X}) = \operatorname{Supp} \mathcal{X}$ and $q(S) = \operatorname{Supp}^{-1} S$. By Theorem 6.6, we have only to show that a thick subcategory $\mathcal{X}$ of $\mathsf{D}^{\mathsf{b}}(R)$ contains the residue field k if and only if $\mathcal{X}$ is a standard thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ not contained in $\mathsf{D}_{\mathsf{perf}}(R)$.

To show the "only if" part, suppose that $\mathcal{X}$ contains k. As R is not regular, k does not belong to $\mathsf{D}_{\mathsf{perf}}(R)$, whence $\mathcal{X}$ is not contained in $\mathsf{D}_{\mathsf{perf}}(R)$. The thick closure $\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\, k$ contains the Koszul complex $\mathrm{K}(\boldsymbol{x}, R)$, where $\boldsymbol{x}$ is a system of generators of the maximal ideal of R. Hence $\mathcal{X}$ contains the nonacyclic perfect complex $\mathrm{K}(\boldsymbol{x}, R)$, which implies that $\mathcal{X}$ is standard.

To show the "if" part, assume that $\mathcal{X}$ is standard and not contained in $\mathsf{D}_{\mathsf{perf}}(R)$. Then the image of $\mathcal{X}$ in $\mathsf{D}_{\mathsf{sg}}(R)$ is nonzero, and hence its thick closure coincides with $\mathsf{D}_{\mathsf{sg}}(R)$ by Proposition 7.7. Therefore $\mathcal{X}$ contains k by Proposition 7.6.

Thus, the proof of the theorem is completed. □

Remark 7.9. Theorem 7.2(2), Proposition 7.7(2) and the statement (written in italic) in Remark 7.4 are valid if one replaces the assumption that *R has minimal multiplicity and k is infinite* with the assumption that *there exists a parameter ideal Q of R with* $\mathfrak{m}^2 = Q\mathfrak{m}$. In fact, the same proofs work under this assumption.

8. Almost Gorenstein rings and Cohen–Macaulay rings of finite CM-representation type

In this section, as another application of Theorem 6.6, we study classifying standard and costandard thick subcategories over an almost Gorenstein ring and a Cohen–Macaulay ring of finite CM-representation type. We start by recalling the definitions:

Definition 8.1. Let R be a Cohen–Macaulay local ring.

(1) We say that R is *almost Gorenstein* if there exists an exact sequence

$$0 \to R \to \omega \to C \to 0$$

of R-modules such that ω is a canonical module of R and C is an *Ulrich* module, that is, C is a Cohen–Macaulay R-module whose multiplicity is equal to the minimal number of generators. For the details of almost Gorenstein local rings, we refer the reader to [Goto et al. 2015].

(2) We say R is of *finite CM-representation type* if there exist only finitely many isomorphism classes of indecomposable maximal Cohen–Macaulay R-modules.

The main result of this section is the following theorem. (The definitions of standard and costandard thick subcategories are given in Section 1.)

Theorem 8.2. *Let R be a non-Gorenstein local ring. Suppose that R is either*

(1) *an almost Gorenstein ring with an isolated singularity*, *or*

(2) *an excellent Cohen–Macaulay ring with canonical module and finite CM-representation type.*

Then there is a one-to-one correspondence

$$\left\{\begin{matrix}\textit{Standard and costandard}\\ \textit{thick subcategories of } \mathsf{D}^{\mathsf{b}}(R)\end{matrix}\right\} \underset{\mathrm{Supp}^{-1}}{\overset{\mathrm{Supp}}{\underset{1-1}{\rightleftarrows}}} \left\{\begin{matrix}\textit{Nonempty specialization-closed}\\ \textit{subsets of } \operatorname{Spec} R\end{matrix}\right\}.$$

We state examples, remarks and propositions related to this theorem, and prove the theorem.

Example 8.3. Let k be an algebraically closed field of characteristic zero. Let t, x, y be indeterminates over k.

(1) Both the numerical semigroup ring $k[[t^3, t^4, t^5]]$ and the Veronese subring $k[[x^3, x^2y, xy^2, y^3]]$ satisfy all the conditions (2) in Theorem 7.2 and (1), (2) in Theorem 8.2.

(2) Consider the numerical semigroup rings $k[[t^4, t^5, t^7]]$, $k[[t^4, t^7, t^9]]$ and the residue ring $k[[x, y, z, s]]/I$, where I is the ideal generated by the 2-minors of the matrix $\left(\begin{smallmatrix} x^2 & y^2-s^{10} & z \\ y & z & x \end{smallmatrix}\right)$. (All of these rings are the completions of positively graded k-algebras.) These rings satisfy Theorem 8.2(1), but do not satisfy Theorem 8.2(2) or Theorem 7.2(2).

For the proofs, see [Goto et al. 2015, Examples 3.16, 7.5, Corollary 11.4 and Theorem 12.1] and [Yoshino 1990, Theorems (9.2) and (10.14)].

In view of this example, it seems that there exist a lot of examples of rings satisfying Theorem 8.2(1). Here is a remark on Theorem 8.2(2).

Remark 8.4. According to [Schreyer 1987, (7.1)], all known examples of a nonhypersurface Cohen–Macaulay complete local $\mathbb{C}$-algebra of finite CM-representation type have minimal multiplicity. Hence, at least for these examples, the one-to-one correspondence in Theorem 8.2 is obtained by restricting that in Theorem 7.2.

The following two propositions play a crucial role in the proof of Theorem 8.2.

Proposition 8.5. *Let R be a local ring with residue field k and dualizing complex D. Assume that k belongs to $\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\{R, D\}$. Let P (resp. I) be a nonacyclic R-complex of finite projective (resp. injective) dimension. Then k belongs to $\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\{P, I\}$.*

Proof. The Foxby equivalence theorem [Avramov and Foxby 1997, Theorem (3.2)] implies that the complex $Q := \mathbf{R}\mathrm{Hom}_R(D, I)$ has finite projective dimension and I is isomorphic to $D \otimes_R^{\boldsymbol{L}} Q$ in $\mathsf{D}^{\mathsf{b}}(R)$. As k is in the thick closure of R and D, applying the functor $- \otimes_R^{\boldsymbol{L}} Q \otimes_R^{\boldsymbol{L}} P$ shows that $k \otimes_R^{\boldsymbol{L}} Q \otimes_R^{\boldsymbol{L}} P$ is in the thick closure of $Q \otimes_R^{\boldsymbol{L}} P$ and $I \otimes_R^{\boldsymbol{L}} P$. Note that $Q \otimes_R^{\boldsymbol{L}} P$ and $I \otimes_R^{\boldsymbol{L}} P$ belong to the thick closures of P and I, respectively. Hence $k \otimes_R^{\boldsymbol{L}} Q \otimes_R^{\boldsymbol{L}} P$ belongs to the thick closure of P and I. Since P and I are not acyclic, $k \otimes_R^{\boldsymbol{L}} Q \otimes_R^{\boldsymbol{L}} P$ is nonzero in $\mathsf{D}^{\mathsf{b}}(R)$, whence it contains $k[n]$ as a direct summand for some integer n. Thus the assertion follows. □

A local ring R is called *G-regular* if the totally reflexive modules over R are the free modules. For the details of G-regular local rings, we refer the reader to [Takahashi 2008].

Proposition 8.6. *Let R be a non-Gorenstein local ring with canonical module ω, being either*

(1) *an almost Gorenstein ring with an isolated singularity*, *or*

(2) *an excellent Cohen–Macaulay ring with canonical module and finite CM-representation type.*

Then $\mathsf{thick}_{\mathsf{mod}\,R}\{R, \omega\} = \mathsf{mod}\,R$. *In particular*, *$R$ is a G-regular local ring.*

Proof. Let k be the residue field of R. We first prove $\mathsf{thick}_{\mathsf{mod}\,R}\{R, \omega\} = \mathsf{mod}\,R$.

(1) Since R is an isolated singularity, we have $\mathsf{thick}_{\mathsf{mod}\,R}\{R, k\} = \mathsf{mod}\,R$ by Corollary 3.5. According to [Goto et al. 2015, Corollary 4.5], there is an exact sequence $0 \to X_n \to \cdots \to X_1 \to X_0 \to k^{r-1} \to 0$, where r is the Cohen–Macaulay type of R and each X_i is a finite direct sum of copies of R and ω. We have $r \geq 2$ since R is not Gorenstein, and it is seen that k belongs to the thick closure of R and ω. Thus the equality follows.

(2) It follows from [Huneke and Leuschke 2002, Corollary 2] that R has an isolated singularity, and so does the completion of R since R is excellent (see [Takahashi 2010, Proposition 3.4]). Using [Takahashi 2013a, Corollary 6.9], one sees that the thick closure of R and ω must be the whole category $\mathsf{mod}\,R$.

Now let us show the last assertion. Let G be a totally reflexive R-module. Let $\mathcal{M}$ be the subcategory of $\mathsf{mod}\,R$ consisting of modules M such that $\mathrm{Ext}_R^{\gg 0}(G, M) = 0$. By definition we have $\mathrm{Ext}_R^{>0}(G, R) = 0$, and moreover $\mathrm{Ext}_R^{>0}(G, \omega) = 0$ since G is maximal Cohen–Macaulay. Therefore R and ω belong to $\mathcal{M}$. It is easy to see that $\mathcal{M}$ is a thick subcategory of $\mathsf{mod}\,R$, whence it contains $\mathsf{thick}_{\mathsf{mod}\,R}\{R, \omega\}$, which coincides with $\mathsf{mod}\,R$. Thus k is in $\mathcal{M}$, which implies that G has finite projective dimension, so that G is free. □

Remark 8.7. In Proposition 8.6(2), the excellence can be replaced with the condition that the completion of R is an isolated singularity.

Now we can give a proof of the main result of this section:

Proof of Theorem 8.2. Let $\mathcal{X}$ be a standard and costandard thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$. Then by Propositions 8.6 and 8.5, $\mathcal{X}$ contains the residue field k of R.

Conversely, let $\mathcal{X}$ be a thick subcategory of $\mathsf{D}^{\mathsf{b}}(R)$ containing k. Then $\mathcal{X}$ contains the Koszul complex $\mathrm{K}(\boldsymbol{x}, R)$ with $\boldsymbol{x}$ a system of generators of the maximal ideal of R, whence $\mathcal{X}$ is standard. By assumption, R admits a canonical module ω. Let $\boldsymbol{y}$ be a system of parameters of R. Then $\boldsymbol{y}$ is a regular sequence on R, and hence on ω. The module $\omega/\boldsymbol{y}\omega$ has finite injective dimension as an R-module, and belongs to $\mathcal{X}$ because it is in $\mathsf{thick}_{\mathsf{D}^{\mathsf{b}}(R)}\, k$. Therefore $\mathcal{X}$ is costandard.

The assertion follows from the above argument and Theorem 6.6. □

9. Gorenstein rings with almost minimal multiplicity

This section is devoted to exploring thick subcategories over a Gorenstein local ring having relatively small multiplicity. Let $(R, \mathfrak{m}, k)$ be a Cohen–Macaulay local ring. We say that R has *almost minimal multiplicity* if the following equality holds:

$$\mathrm{e}(R) = \mathrm{edim}\, R - \dim R + 2.$$

Assume that k is infinite. Then there is a minimal reduction Q of $\mathfrak{m}$. Note that Q is a parameter ideal of R satisfying $\mathfrak{m}^2/Q\mathfrak{m} \cong k$, and hence $\mathfrak{m}^3 \subseteq Q$. Only assuming this inclusion, we have the following classification of thick subcategories, which is the main result of this section:

Theorem 9.1. *Let $(R, \mathfrak{m}, k)$ be a Gorenstein nonregular local ring with an isolated singularity. Let Q be a parameter ideal of R containing $\mathfrak{m}^3$. Then there is a one-to-one correspondence*

$$\left\{\begin{matrix} \textit{Thick subcategories} \\ \textit{of } \mathsf{D}^{\mathsf{b}}(R) \textit{ containing} \\ \textit{a nonacyclic perfect } \bar{R}\textit{-complex} \end{matrix}\right\} \underset{\mathrm{Supp}^{-1}}{\overset{\mathrm{Supp}}{\underset{1-1}{\rightleftarrows}}} \left\{\begin{matrix} \textit{Nonempty specialization-closed} \\ \textit{subsets of } \mathrm{Spec}\, R \end{matrix}\right\},$$

where $\bar{R} = R/(Q : \mathfrak{m})$.

We first prepare lemmas to show this theorem.

Lemma 9.2. *Let* $(R, \mathfrak{m}, k)$ *be an artinian Gorenstein local ring which is not a field. Then* $\mathsf{thick}_{\mathsf{mod}\,R}(R/\operatorname{Soc} R) = \mathsf{mod}\,R$.

Proof. Denote by $(-)^*$ the R-dual functor $\operatorname{Hom}_R(-, R)$. Since R is artinian and Gorenstein, R is an injective R-module and $k^* \cong k$. Applying $(-)^*$ to the natural exact sequence $0 \to \mathfrak{m} \to R \to k \to 0$, we have an exact sequence

$$0 \to k \to R \to R/\operatorname{Soc} R \to 0 \tag{9.2.1}$$

and see that $\mathfrak{m}^*$ is isomorphic to $R/\operatorname{Soc} R$.

Let $x_1, x_2, \dots, x_n$ be a minimal system of generators of $\mathfrak{m}$. As R is not a field, the integer n is positive. Let $I = (x_1^2, x_2, \dots, x_n)$ be an ideal. Then $\mathfrak{m}/I$ is isomorphic to k, and there exists an exact sequence $0 \to k \to R/I \to k \to 0$. Taking the first syzygies, we obtain an exact sequence $0 \to \mathfrak{m} \to R \oplus I \to \mathfrak{m} \to 0$. Applying $(-)^*$ gives rise to an exact sequence $0 \to \mathfrak{m}^* \to R \oplus I^* \to \mathfrak{m}^* \to 0$.

Thus, there is an exact sequence

$$0 \to R/\operatorname{Soc} R \to R \oplus I^* \to R/\operatorname{Soc} R \to 0. \tag{9.2.2}$$

It follows from (9.2.2) that $\mathsf{thick}_{\mathsf{mod}\,R}(R/\operatorname{Soc} R)$ contains R, and from (9.2.1) that it contains k. Since R is artinian, $\mathsf{thick}_{\mathsf{mod}\,R}(R/\operatorname{Soc} R)$ coincides with the whole module category $\mathsf{mod}\,R$. ☐

We need one more lemma, whose proof is straightforward.

Lemma 9.3. *Let* $\mathcal{T}, \mathcal{U}$ *be triangulated categories, and let* $F : \mathcal{T} \to \mathcal{U}$ *be a triangle functor. Let* $\mathcal{X}$ *be a thick subcategory of* $\mathcal{U}$. *Denote by* $F^{-1}(\mathcal{X})$ *the subcategory of* $\mathcal{T}$ *consisting of objects* $T \in \mathcal{T}$ *with* $F(T) \in \mathcal{X}$. *Then* $F^{-1}(\mathcal{X})$ *is a thick subcategory of* $\mathcal{T}$.

Now we can prove our Theorem 9.1:

Proof of Theorem 9.1. By Theorem 6.6, it suffices to show that a thick subcategory $\mathcal{X}$ of $\mathsf{D}^{\mathsf{b}}(R)$ contains a nonacyclic perfect $\bar{R}$-complex if and only if $\mathcal{X}$ contains the residue field $k = R/\mathfrak{m}$.

The "if" part: If k is in $\mathcal{X}$, then all R-modules of finite length are in $\mathcal{X}$, whence $\bar{R} \in \mathcal{X}$.

The "only if" part: Assume that $\mathcal{X}$ contains a nonacyclic perfect $\bar{R}$-complex L. Let $F : \mathsf{D}^{\mathsf{b}}(\bar{R}) \to \mathsf{D}^{\mathsf{b}}(R)$ be the natural triangle functor. Lemma 9.3 implies that $F^{-1}(\mathcal{X})$ is a thick subcategory of $\mathsf{D}^{\mathsf{b}}(\bar{R})$, and it is standard since it contains L. As $\mathfrak{m}^3$ is contained in Q, the square of the maximal ideal of $\bar{R}$ is zero. Using Remarks 7.4 and 7.9, we observe that $F^{-1}(\mathcal{X})$ either contains k or is contained in $\mathsf{D}_{\mathsf{perf}}(\bar{R})$. As to the former case, k belongs to $\mathcal{X}$.

Let us consider the latter case. Note that $F^{-1}(\mathcal{X})$ is a thick subcategory of $\mathsf{D}_{\mathsf{perf}}(\bar{R})$, and that $\operatorname{Spec}\bar{R}$ consists of the maximal ideal. By [Neeman 1992, Theorem 1.5], $F^{-1}(\mathcal{X})$ coincides with either the zero category $\mathbf{0}$ or the whole category $\mathsf{D}_{\mathsf{perf}}(\bar{R})$. Because the nonacyclic complex L is in $F^{-1}(\mathcal{X})$, we have $F^{-1}(\mathcal{X}) = \mathsf{D}_{\mathsf{perf}}(\bar{R})$. In particular, $\mathcal{X}$ contains $\bar{R}$. Note that R/Q is an artinian Gorenstein ring that is not a field. Applying Lemma 9.2 to the ring R/Q, we have $\mathsf{thick}_{\mathsf{mod}\,R/Q}(\bar{R}) = \mathsf{mod}\,R/Q$. Hence k is in $\mathsf{thick}_{\mathsf{D^b}(R/Q)}(\bar{R})$. Sending this containment by the natural triangle functor $\mathsf{D^b}(R/Q) \to \mathsf{D^b}(R)$ shows $k \in \mathsf{thick}_{\mathsf{D^b}(R)}(\bar{R})$. Thus k belongs to $\mathcal{X}$. □

Acknowledgments

Part of this work was done during the visit of the author to the Centre de Recerca Matemàtica (CRM) to attend the research program "IRTATCA: Interactions between Representation Theory, Algebraic Topology and Commutative Algebra" held from January to June in 2015. The author is grateful for their great atmosphere and kind hospitality. Also, the author very much thanks Mitsuyasu Hashimoto for giving helpful advice on the proof of Lemma 3.2, Shunsuke Takagi for indicating the reference [Huneke and Watanabe 2015], an anonymous referee for valuable comments on the proof of Theorem 6.6 (see Remark 6.7), and another anonymous referee for a lot of useful comments and suggestions throughout the paper.

References

[Auslander and Bridger 1969] M. Auslander and M. Bridger, *Stable module theory*, Memoirs of the American Mathematical Society **94**, American Mathematical Society, Providence, RI, 1969. MR Zbl

[Avramov and Foxby 1997] L. L. Avramov and H.-B. Foxby, "Ring homomorphisms and finite Gorenstein dimension", *Proc. London Math. Soc.* (3) **75**:2 (1997), 241–270. MR Zbl

[Bruns and Herzog 1998] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, 2nd ed., Cambridge Studies in Advanced Mathematics **39**, Cambridge Univ. Press, 1998. Zbl

[Dwyer et al. 2006] W. Dwyer, J. P. C. Greenlees, and S. Iyengar, "Finiteness in derived categories of local rings", *Comment. Math. Helv.* **81**:2 (2006), 383–432. MR Zbl

[Gabriel 1962] P. Gabriel, "Des catégories abéliennes", *Bull. Soc. Math. France* **90** (1962), 323–448. MR Zbl

[Goto et al. 2015] S. Goto, R. Takahashi, and N. Taniguchi, "Almost Gorenstein rings: towards a theory of higher dimension", *J. Pure Appl. Algebra* **219**:7 (2015), 2666–2712. MR Zbl

[Huneke and Leuschke 2002] C. Huneke and G. J. Leuschke, "Two theorems about maximal Cohen–Macaulay modules", *Math. Ann.* **324**:2 (2002), 391–404. MR Zbl

[Huneke and Watanabe 2015] C. Huneke and K.-i. Watanabe, "Upper bound of multiplicity of F-pure rings", *Proc. Amer. Math. Soc.* **143**:12 (2015), 5021–5026. MR Zbl

[Iyengar 2004] S. Iyengar, "Modules and cohomology over group algebras: one commutative algebraist's perspective", pp. 51–86 in *Trends in commutative algebra*, edited by L. L. Avramov et al., Math. Sci. Res. Inst. Publ. **51**, Cambridge Univ. Press, 2004. MR Zbl

[Krause and Stevenson 2013] H. Krause and G. Stevenson, "A note on thick subcategories of stable derived categories", *Nagoya Math. J.* **212** (2013), 87–96. MR Zbl

[Neeman 1992] A. Neeman, "The chromatic tower for $D(R)$", *Topology* **31**:3 (1992), 519–532. MR Zbl

[Orlov 2011] D. Orlov, "Formal completions and idempotent completions of triangulated categories of singularities", *Adv. Math.* **226**:1 (2011), 206–217. MR Zbl

[Schoutens 2003] H. Schoutens, "Projective dimension and the singular locus", *Comm. Algebra* **31**:1 (2003), 217–239. MR Zbl

[Schreyer 1987] F.-O. Schreyer, "Finite and countable CM-representation type", pp. 9–34 in *Singularities, representation of algebras, and vector bundles* (Lambrecht, 1985), edited by G.-M. Greuel and G. Trautmann, Lecture Notes in Math. **1273**, Springer, Berlin, 1987. MR Zbl

[Stevenson 2014] G. Stevenson, "Duality for bounded derived categories of complete intersections", *Bull. Lond. Math. Soc.* **46**:2 (2014), 245–257. MR Zbl

[Takahashi 2008] R. Takahashi, "On G-regular local rings", *Comm. Algebra* **36**:12 (2008), 4472–4491. MR Zbl

[Takahashi 2010] R. Takahashi, "Classifying thick subcategories of the stable category of Cohen–Macaulay modules", *Adv. Math.* **225**:4 (2010), 2076–2116. MR Zbl

[Takahashi 2013a] R. Takahashi, "Classifying resolving subcategories over a Cohen–Macaulay local ring", *Math. Z.* **273**:1-2 (2013), 569–587. MR Zbl

[Takahashi 2013b] R. Takahashi, "Thick subcategories over Gorenstein local rings that are locally hypersurfaces on the punctured spectra", *J. Math. Soc. Japan* **65**:2 (2013), 357–374. MR Zbl

[Thomason 1997] R. W. Thomason, "The classification of triangulated subcategories", *Compositio Math.* **105**:1 (1997), 1–27. MR Zbl

[Yoshino 1990] Y. Yoshino, *Cohen–Macaulay modules over Cohen–Macaulay rings*, London Mathematical Society Lecture Note Series **146**, Cambridge Univ. Press, 1990. MR Zbl

Received May 15, 2016. Revised March 28, 2017.

RYO TAKAHASHI
GRADUATE SCHOOL OF MATHEMATICS
NAGOYA UNIVERSITY
NAGOYA
JAPAN

takahashi@math.nagoya-u.ac.jp
http://www.math.nagoya-u.ac.jp/~takahashi/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.213

PROJECTIONS IN THE CURVE COMPLEX ARISING FROM COVERING MAPS

ROBERT TANG

Let $P : \Sigma \to S$ be a finite degree covering map between surfaces. Rafi and Schleimer showed that there is an induced quasi-isometric embedding $\Pi : \mathcal{C}(S) \to \mathcal{C}(\Sigma)$ between the associated curve complexes. We define an operation on curves in $\mathcal{C}(\Sigma)$ using minimal intersection number conditions and prove that it approximates a nearest point projection to $\Pi(\mathcal{C}(S))$. We also approximate hulls of finite sets of vertices in the curve complex, together with their corresponding nearest point projections, using intersection numbers.

1. Overview

Let S be a closed, orientable, connected surface of genus $g \geq 0$ with $m \geq 0$ marked points whose *complexity* $\xi(S) := 3g - 3 + m$ is at least 2. The *curve complex* of S, denoted $\mathcal{C}(S)$, is the simplicial complex whose vertices are free isotopy classes of simple closed curves on S and whose simplices are spanned by multicurves. The curve complex has seen much activity in recent years due to its connections to mapping class groups, Teichmüller theory, and the geometry of hyperbolic 3-manifolds.

Given a finite degree covering map $P : \Sigma \to S$ and a simple closed curve $a \in \mathcal{C}(S)$, the preimage $P^{-1}(a)$ is a disjoint union of simple closed curves on Σ. Rafi and Schleimer, using techniques from Teichmüller theory, proved that the (one-to-many) *lifting operation* $\Pi : \mathcal{C}(S) \to \mathcal{C}(\Sigma)$ defined by setting $\Pi(a) = P^{-1}(a)$ is a quasi-isometric embedding. In [Tang 2012], we give a new proof using results from hyperbolic 3-manifold geometry.

Theorem 1.1 [Rafi and Schleimer 2009]. *Let $P : \Sigma \to S$ be a finite degree covering map. Then the map $\Pi : \mathcal{C}(S) \to \mathcal{C}(\Sigma)$ defined above is a Λ-quasi-isometric embedding*, *where Λ depends only on $\xi(\Sigma)$ and* deg P.

The primary aim of this paper is to give a combinatorial approximation of the nearest point projection map to the image of Π. We define an operation $\pi : \mathcal{C}(\Sigma) \to \Pi(\mathcal{C}(S)) \subseteq \mathcal{C}(\Sigma)$ as follows: Given a curve $\alpha \in \mathcal{C}(\Sigma)$, let $\pi(\alpha) = \Pi(b)$

MSC2010: 20F65, 57M50.
Keywords: curve complex, covering map, nearest point projection, hull.

where b is a curve which has minimal intersection number with $P(\alpha)$ among all curves in $\mathcal{C}(S)$.

Theorem 6.1. *Let $P : \Sigma \to S$ be a finite degree covering map and suppose $\alpha \in \mathcal{C}(\Sigma)$ is a curve. Then $\pi(\alpha)$ is a uniformly bounded distance from any nearest point projection of α to $\Pi(\mathcal{C}(S))$ in $\mathcal{C}(\Sigma)$, where the bounds depend only on $\xi(\Sigma)$ and the degree of P.*

Proposition 6.2. *Assume further that P is regular, with deck group G. Then $\pi(\alpha)$ is a uniformly bounded distance from any circumcenter for the G-orbit of α in $\mathcal{C}(\Sigma)$. Moreover, the bounds depend only on $\xi(\Sigma)$ and the degree of P.*

The main tools we develop in order to prove our main results are descriptions of hulls in the curve complex using intersection number conditions, which may be of independent interest. These generalize Bowditch's [2006b] approximation of quasigeodesics in $\mathcal{C}(S)$ using intersection numbers, and Masur and Minsky's [1999] notion of balance time for a curve on a Teichmüller geodesic. Our results rely on the geometry of singular Euclidean surfaces used to estimate weighted intersection numbers. We state simplified versions of the relevant propositions below — see Section 5 for more precise formulations.

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ be an n-tuple of distinct curves in $\mathcal{C}(S)$, where $n \geq 2$. Given a nonzero vector $\mathbf{t} = (t_1, \ldots, t_n)$ of nonnegative reals, let $\gamma_{\mathbf{t}} \in \mathcal{C}(S)$ be a curve which minimizes the weighted intersection number $\sum_i t_i i(\alpha_i, \cdot)$. Define the *hyperbolic hull* $\mathrm{Hull}(\boldsymbol{\alpha})$ to be the union of all geodesic segments in $\mathcal{C}(S)$ connecting a pair of points in $\boldsymbol{\alpha}$ (viewed as a vertex set in $\mathcal{C}(S)$).

Proposition 5.2. *The sets $\mathrm{Hull}(\boldsymbol{\alpha})$ and $\bigcup_{\mathbf{t}} \gamma_{\mathbf{t}}$ agree up to a uniformly bounded Hausdorff distance in $\mathcal{C}(S)$, where the union is taken over all nonzero $\mathbf{t} \in \mathbb{R}^n_{\geq 0}$. Moreover, the bound depends only on $\xi(S)$ and n.*

Proposition 5.5. *Suppose $\beta \in \mathcal{C}(S)$ is a curve satisfying $i(\alpha_i, \beta) \neq 0$ for all i. Let the **balance vector** $\mathbf{t}_\beta = (t_1, \ldots, t_n)$ of β with respect to $\boldsymbol{\alpha}$ be given by $t_i = i(\alpha_i, \beta)^{-1}$ for each i. Then $\gamma_{\mathbf{t}_\beta}$ is a uniformly bounded distance from any nearest point projection of β to $\mathrm{Hull}(\boldsymbol{\alpha})$ in $\mathcal{C}(S)$, where the bound depends only on $\xi(S)$ and n.*

Organization. We review the curve complex in Section 2, and some coarse geometric notions in Section 3, placing a particular emphasis on δ-hyperbolic spaces.

In Section 4, we introduce a generalization of Bowditch's [2006b] construction of singular Euclidean structures on surfaces on which the geodesic lengths of curves estimate suitable weighted intersection numbers. We verify in Section 7 that these surfaces satisfy a quadratic isoperimetric inequality and then apply a theorem of Bowditch to establish the existence of wide annuli.

In Section 5, we introduce two notions of hulls for finite sets in $\mathcal{C}(S)$: one arising geometrically in $\mathcal{C}(S)$; the other defined using intersection number conditions. We give proofs of Propositions 5.2 and 5.5 assuming bounded diameter properties for sets of curves satisfying certain bounded weighted intersection number conditions (Lemma 5.1) — a key fact whose proof we defer to Section 8. In Section 6, we utilize the results from Sections 3 and 5 to give proofs of the main theorems.

2. The curve complex

Let $S = (S, \Omega)$ denote a closed, orientable, connected surface of genus $g \geq 0$ together with a set Ω of $m \geq 0$ marked points. A *curve* on S is a continuous map $a : S^1 \to S - \Omega$. We will also write a for its image on S. A curve a is *simple* if it is an embedded copy of S^1. We call a curve *trivial* or *peripheral* if it is freely homotopic to a curve bounding a disc or a disc with exactly one marked point, respectively. A simple closed curve which is nontrivial and nonperipheral is called *essential*. A *multicurve* on S is a finite collection of nonparallel essential simple closed curves which can be realized disjointly simultaneously.

Let $\mathcal{C}^0(S)$ denote the set of free homotopy classes of essential simple closed curves on S. Unless explicitly stated otherwise, we will blur the distinction between curves and their free homotopy classes. In this paper, we assume that S has *complexity* $\xi(S) := 3g - 3 + m$ at least 2; modifications to the following definition are required for low-complexity cases but we shall not deal with them here. For an introduction to the curve complex, see [Schleimer 2005].

Definition 2.1. The *curve complex* of S, denoted $\mathcal{C}(S)$, is a simplicial complex whose vertex set is $\mathcal{C}^0(S)$ and whose simplices are spanned by multicurves. In particular, two distinct simple closed curves are connected by an edge in $\mathcal{C}(S)$ if and only if they have disjoint representatives on S.

For our purposes, it suffices to study the 1-skeleton $\mathcal{C}^1(S)$ of the curve complex, known as the *curve graph*. Indeed $\mathcal{C}^1(S)$ equipped with the induced path metric, denoted d_S, is naturally *quasi-isometric* to $\mathcal{C}(S)$ with the standard simplicial metric. To simplify notation, we shall write $\mathcal{C}(S)$ in place of $\mathcal{C}^1(S)$ and $\alpha \in \mathcal{C}(S)$ to denote a curve (or multicurve).

A finite collection of curves *fills* S if their complement is a disjoint union of discs each with at most one marked point. Note that a pair $\alpha, \beta \in \mathcal{C}(S)$ fill S if and only if $d_S(\alpha, \beta) \geq 3$. Given free homotopy classes of curves α and β, not necessarily simple, define their *(geometric) intersection number* $i(\alpha, \beta)$ to be the minimal value of $|a \cap b|$ over all representatives $a \in \alpha$ and $b \in \beta$ in general position on S.

Lemma 2.2 [Hempel 2001], [Schleimer 2005]. *Suppose α and β are curves in $\mathcal{C}(S)$. Then*

$$d_S(\alpha, \beta) \leq 2 \log_2 i(\alpha, \beta) + 2$$

whenever $i(\alpha, \beta) \neq 0$.

As an immediate corollary, we see that $\mathcal{C}(S)$ is connected (this was originally observed by Harvey [1981]). The curve graph is also locally infinite and has infinite diameter [Kobayashi 1988]. Masur and Minsky [1999] proved the following celebrated theorem regarding the large scale geometry of the curve graph:

Theorem 2.3. *Given any surface S with $\xi(S) \geq 2$, there exists $\delta > 0$ so that the curve graph $\mathcal{C}(S)$ is δ-hyperbolic.*

Bowditch [2006b] gives a combinatorial proof of hyperbolicity using intersection numbers. We will be extending many of the results established in his paper in Sections 4 and 5.

Theorem 2.4 [Bowditch 2014], [Aougab 2013], [Clay et al. 2014], [Hensel et al. 2015]. *The constant $\delta > 0$ in Theorem 2.3 can be chosen independently of S.*

Hensel, Przytycki and Webb in particular show that all geodesic triangles in $\mathcal{C}(S)$ possess 17-centers.

3. Coarse geometry

We now recall some basic notions concerning Gromov hyperbolic spaces. Most of the statements and results are either well known in the literature or are relatively straightforward to deduce. We refer the reader to [Bridson and Haefliger 1999], [Gromov 1987], [Alonso et al. 1991], and [Bowditch 2006a] for more background, and to [Tang 2013] for most of the proofs.

3A. ***Notation.*** Let $(\mathcal{X}, d)$ be a metric space. Given any subset $A \subseteq \mathcal{X}$ and a point $x \in \mathcal{X}$, we define $d(x, A) := \inf\{d(x, a) \mid a \in A\}$. For $\mathsf{r} \geq 0$, let

$$\mathcal{N}_{\mathsf{r}}(A) = \{x \in \mathcal{X} \mid d(x, A) \leq \mathsf{r}\}$$

denote the r-*neighborhood* of A in $\mathcal{X}$. For subsets $A, B \subseteq \mathcal{X}$ and $\mathsf{r} \geq 0$, write

$$A \subseteq_{\mathsf{r}} B \iff A \subseteq \mathcal{N}_{\mathsf{r}}(B)$$

and

$$A \approx_{\mathsf{r}} B \iff A \subseteq_{\mathsf{r}} B \text{ and } B \subseteq_{\mathsf{r}} A.$$

Define the *Hausdorff distance* between A and B to be

$$\operatorname{HausDist}(A, B) = \inf\{\mathsf{r} \geq 0 \mid A \approx_{\mathsf{r}} B\}.$$

To simplify notation, we will often write $a \in \mathcal{X}$ in place of a singleton set $\{a\} \subseteq \mathcal{X}$. If a and b are real numbers then write

$$a \approx_{\mathsf{r}} b \iff |a - b| \leq \mathsf{r}.$$

The *diameter* of $A \subseteq \mathcal{X}$ is defined to be

$$\operatorname{diam}(A) := \sup\{d(x, y) \mid x, y \in A\}.$$

We will also abbreviate $d(x, y)$ to xy if there is no chance of confusion.

3B. ***Geodesics, quasiconvexity and quasi-isometries.*** Let $I \subseteq \mathbb{R}$ be an interval. A *geodesic* is a map $\gamma : I \to \mathcal{X}$ so that $d(\gamma(t), \gamma(s)) = |t - s|$ for all $t, s \in I$. A *geodesic segment* connecting points x and y in $\mathcal{X}$ is the image of a geodesic $\gamma : [0, d(x, y)] \to \mathcal{X}$ such that $\gamma(0) = x$ and $\gamma(d(x, y)) = y$. A metric space $\mathcal{X}$ is called a *geodesic space* if every pair of points can be connected by a geodesic segment. A subset $U \subseteq \mathcal{X}$ is Q-*quasiconvex* if every geodesic segment connecting any pair of points in U lies in $\mathcal{N}_{\mathsf{Q}}(U)$. We say a subset is *quasiconvex* if it is Q-quasiconvex for some $\mathsf{Q} \geq 0$.

A (one-to-many) map $f : \mathcal{X} \to \mathcal{Y}$ between metric spaces is a Λ-*quasi-isometric embedding* if for all $x_1, x_2 \in \mathcal{X}$ and $y_1 \in f(x_1),\ y_2 \in f(x_2)$ we have

$$d_{\mathcal{Y}}(y_1, y_2) \leq \Lambda d_{\mathcal{X}}(x_1, x_2) + \Lambda \quad \text{and} \quad d_{\mathcal{X}}(x_1, x_2) \leq \Lambda d_{\mathcal{Y}}(y_1, y_2) + \Lambda.$$

In addition, if $\mathcal{N}_{\Lambda}(f(\mathcal{X})) = \mathcal{Y}$ then f is called a Λ-*quasi-isometry* and we say that $\mathcal{X}$ and $\mathcal{Y}$ are Λ-*quasi-isometric*. If $\mathcal{X}$ and $\mathcal{Y}$ are Λ-quasi-isometric for some $\Lambda \geq 1$ then we may simply say that they are *quasi-isometric*.

3C. ***Gromov hyperbolic spaces.*** Throughout this paper, we shall use the *thin triangles* definition of δ-hyperbolicity which we now describe.

Let $(\mathcal{X}, d)$ be a geodesic space. Let $T = [x, y] \cup [y, z] \cup [z, x]$ be a geodesic triangle in $\mathcal{X}$ with corners at $x, y, z \in \mathcal{X}$. There exist unique *internal points* $o_x \in [y, z],\ o_y \in [x, z]$, and $o_z \in [x, y]$ such that $xo_y = xo_z,\ yo_x = yo_z$, and $zo_x = zo_y$. The internal points cut T into three pairs of geodesic segments; each pair consists of two segments of equal length emanating from the same corner of T. We say that T is δ-*thin* if each pair of segments δ-fellow travel: for all $u \in [x, o_y]$ and $v \in [x, o_z]$ satisfying $xu = xv$, we have $uv \leq \delta$ (and similarly for the other two pairs). We say $\mathcal{X}$ is δ-*hyperbolic* if every geodesic triangle in $\mathcal{X}$ is δ-thin.

We will also use some equivalent notions of Gromov hyperbolicity:

Lemma 3.1 (four point condition, [Bridson and Haefliger 1999] Proposition 1.22). *Let $\mathcal{X}$ be a geodesic space. If $\mathcal{X}$ is δ-hyperbolic then*

$$xy + zw \leq \max\{xz + yw, xw + yz\} + 2\delta$$

for all $x, y, z, w \in \mathcal{X}$. Conversely, if this inequality holds for all points x, y, z and w in $\mathcal{X}$, then $\mathcal{X}$ is δ'-hyperbolic for some $\delta' \geq 0$ depending only on δ.

Suppose $\mathsf{k} \geq 0$. A k-*center* for a geodesic triangle $T \subseteq \mathcal{X}$ is a point in $\mathcal{X}$ which lies within a distance k of each side of T.

Lemma 3.2 [Bowditch 2006a, Proposition 6.13]. *Any geodesic triangle in a δ-hyperbolic space possesses a δ-center, namely, any of its internal points. Conversely, if $\mathcal{X}$ is a geodesic space for which there is some $\mathsf{k} \geq 0$ such that all geodesic triangles in $\mathcal{X}$ possess k-centers then $\mathcal{X}$ is δ-hyperbolic for some $\delta \geq 0$ depending only on k.*

3D. ***Nearest point projections to quasiconvex sets.*** Let $\mathcal{X}$ be a δ-hyperbolic space, and $U \subseteq \mathcal{X}$ be a closed, nonempty Q-quasiconvex subset. The set of *nearest point projections* of a point $x \in \mathcal{X}$ to U in $\mathcal{X}$ is

$$\mathrm{proj}_U(x) := \{p \in U \mid xp = d(x, U)\}.$$

Since U is closed, $\mathrm{proj}_U(x)$ is nonempty.

Lemma 3.3. *For all $x \in \mathcal{X}$, we have* $\mathrm{diam}(\mathrm{proj}_U(x)) \leq 2\delta + 2\mathsf{Q}$.

Lemma 3.4. *Given $x \in \mathcal{X}$, let $p \in \mathrm{proj}_U(x)$ be any nearest point projection. Then for all $u \in U$, $[x, u] \approx_{2\delta+\mathsf{Q}} [x, p] \cup [p, u]$ and $xu \approx_{2\delta+2\mathsf{Q}} xp + pu$.*

For $\mathsf{r} \geq 0$, call $q \in U$ an r-*entry point* of x to U if for every $u \in U$, all geodesics from x to u intersect $\mathcal{N}_{\mathsf{r}}(q)$. Let $\mathrm{entry}_U(x, \mathsf{r})$ denote the set of such points.

Lemma 3.5. *Let $\mathsf{r} \geq 0$. Then for all $x \in \mathcal{X}$, we have* $\mathrm{entry}_U(x, \mathsf{r}) \subseteq_{2\mathsf{r}} \mathrm{proj}_U(x)$. *Furthermore, if $\mathsf{r} \geq 2\delta + \mathsf{Q}$ then* $\mathrm{entry}_U(x, \mathsf{r}) \approx_{2\mathsf{r}} \mathrm{proj}_U(x)$.

We shall also need the fact that nearest point projections to quasiconvex sets are well behaved under quasi-isometric embeddings.

Lemma 3.6. *Let $f : \mathcal{X} \to \mathcal{X}'$ be a Λ-quasi-isometric embedding of geodesic spaces, where $\mathcal{X}'$ is δ'-hyperbolic. Let C be a Q-quasiconvex subset of $\mathcal{X}$ and let $C' = f(C)$. Given a point $x \in \mathcal{X}$, let x' be a point in $f(x)$. Let p and q' be nearest point projections of x to C and x' to C' respectively. Let $q \in \mathcal{X}$ be a point satisfying $q' \in f(q)$. Then $p \approx_{\mathsf{K}} q$, where K depends only on δ', Λ and Q.*

Proof. First, note that $\mathcal{X}$ is δ-hyperbolic and C' is Q'-quasiconvex in $\mathcal{X}'$ for some constants $\delta = \delta(\Lambda, \delta')$ and $\mathsf{Q}' = \mathsf{Q}'(\mathsf{Q}, \Lambda, \delta)$. Let $c \in \mathcal{X}$ be a k-center for x, p and q, where $\mathsf{k} = \delta$. Any point $c' \in f(c)$ is then a k'-center for x', p' and q', where $\mathsf{k}' = \mathsf{k}'(\mathsf{k}, \Lambda)$ and $p' \in f(p)$. One can check that $xp \approx_{2\mathsf{k}} xc + cp$. By quasiconvexity of C, there is some point $y \in C$ satisfying $cy \leq \mathsf{k} + \mathsf{Q}$. Since p is a nearest point projection of x to C, we obtain

$$xc + cp - 2\mathsf{k} \leq xp \leq xy \leq xc + cy \leq xc + \mathsf{k} + \mathsf{Q},$$

which implies $cp \leq \mathsf{Q} + 3\mathsf{k}$. Similarly, we can deduce $c'q' \leq \mathsf{Q}' + 3\mathsf{k}'$. Since f is a Λ-quasi-isometric embedding, it follows that $cq \leq \Lambda \times c'q' + \Lambda$ and hence

$$pq \leq pc + cq \leq \mathsf{K},$$

where $\mathsf{K} = \mathsf{Q} + 3\mathsf{k} + \Lambda(\mathsf{Q}' + 3\mathsf{k}') + \Lambda$. □

3E. ***Hyperbolic hulls.*** Let $\mathcal{X}$ be a δ-hyperbolic space and suppose $U \subseteq \mathcal{X}$ is nonempty. The *hyperbolic hull* of U, denoted $\mathrm{Hull}(U)$, is the union of all geodesic segments in $\mathcal{X}$ connecting a pair of points in U.

Example 3.7. Let U be a finite subset of $\mathbb{H}^n$, where $n \geq 1$. Then $\mathrm{Hull}(U)$ is a uniformly bounded Hausdorff distance away from the convex hull of U in $\mathbb{H}^n$.

Lemma 3.8. *The hyperbolic hull* $\mathrm{Hull}(U)$ *is* 2δ*-quasiconvex. Furthermore, if* $C \subseteq \mathcal{X}$ *is a* Q*-quasiconvex set which contains* U *then* $\mathrm{Hull}(U) \subseteq_{\mathsf{Q}} C$.

In fact, these properties characterize $\mathrm{Hull}(U)$ up to finite Hausdorff distance.

Corollary 3.9. *Let* $C \subseteq \mathcal{X}$ *be a* Q*-quasiconvex set containing* U *with the following property*: *for any* Q'*-quasiconvex set* $C' \subseteq \mathcal{X}$ *also containing* U, *we have* $C \subseteq_{\mathsf{r}} C'$ *for some* $\mathsf{r} = \mathsf{r}(\mathsf{Q}, \mathsf{Q}') \geq 0$. *Then* $\mathrm{HausDist}(C, \mathrm{Hull}(U)) \leq \max\{\mathsf{Q}, \mathsf{r}(\mathsf{Q}, 2\delta)\}$.

3F. ***Circumcenters.*** Let U be a nonempty finite subset of a δ-hyperbolic space $\mathcal{X}$. The *radius* of U is

$$\mathrm{rad}(U) := \min\{\mathsf{r} \geq 0 \mid \text{there exists } x \in \mathcal{X}, U \subseteq \mathcal{N}_{\mathsf{r}}(x)\}.$$

Call $x \in \mathcal{X}$ a *circumcenter* of U if $U \subseteq \mathcal{N}_{\mathsf{r}}(x)$, where $\mathsf{r} = \mathrm{rad}(U)$, and write $\mathrm{circ}(U)$ for the set of circumcenters of U.

Lemma 3.10. *Suppose* $x \in \mathcal{X}$ *satisfies* $U \subseteq \mathcal{N}_{\mathsf{r}+\epsilon}(x)$, *where* $\mathsf{r} = \mathrm{rad}(U)$ *and* $\epsilon \geq 0$. *Then for any* $c \in \mathrm{circ}(U)$, *we have* $cx \leq 2\delta + 2\epsilon$ *and hence* $\mathrm{diam}(\mathrm{circ}(U)) \leq 2\delta$.

Lemma 3.11. *Let* $x, y \in U$ *be points such that* $xy \geq \mathrm{diam}(U) - 2\epsilon$, *for some* $\epsilon \geq 0$. *Let* m *be the midpoint of a geodesic segment* $[x, y]$. *Then* $c \approx_{2\delta+\epsilon} m$, *where* c *is any circumcenter of* U. *Furthermore, we have* $\mathrm{diam}(U) \leq 2\,\mathrm{rad}(U) \leq \mathrm{diam}(U) + 2\delta$.

We also give the following characterization of circumcenters of orbits under finite group actions on δ-hyperbolic spaces:

Lemma 3.12. *Assume* G *is a finite group acting by isometries on a* δ*-hyperbolic space* $\mathcal{X}$. *Fix a point* $x_0 \in \mathcal{X}$ *and let* c *be a circumcenter for* Gx_0. *Given a point* $z \in \mathcal{X}$, *let* p *be any of its nearest point projection to* $\mathrm{Hull}(Gx_0)$. *Then*

$$pc \leq \mathrm{rad}(Gz) + 7\delta$$

and hence

$$zc \leq \mathrm{rad}(Gz) + d(z, \mathrm{Hull}(Gx_0)) + 7\delta.$$

Proof. We first claim that p lies within a distance δ of a geodesic segment $[u, v]$, where $u, v \in Gx_0$ are points such that $uv \geq \mathrm{diam}(Gx_0) - 2\delta$. Suppose p lies on a geodesic segment $[x, y]$ for some x and y in Gx_0. There exist some x' and y' in Gx_0 such that $xx' = yy' = \mathrm{diam}(Gx_0)$. If $x' = y'$ then the claim follows from hyperbolicity. Now assume $x' \neq y'$. By Lemma 3.1, we have

$$2\,\mathrm{diam}(Gx_0) = xx' + yy' \geq \max\{xy + x'y', xy' + x'y\} \geq 2\,\mathrm{diam}(Gx_0) - 2\delta.$$

If $xy + x'y' \geq 2\,\mathrm{diam}(Gx_0) - 2\delta$, then $xy \geq \mathrm{diam}(Gx_0) - 2\delta$, which implies the claim. If not, then $xy' \geq \mathrm{diam}(Gx_0) - 2\delta$. The claim then follows by considering a geodesic triangle with x, y and y' as its vertices.

Now suppose $q \in [u, v]$ is a point such that $pq \leq \delta$. Then

$$d(z, [u, v]) \leq zq \leq zp + pq \leq d(z, \mathrm{Hull}(Gx_0)) + \delta \leq d(z, [u, v]) + \delta.$$

By considering a geodesic triangle with vertices u, v, and z, one can show that $q \approx_{3\delta} o$, where $o \in [u, v]$ is the internal point opposite z. Observe that

$$d(z, x_0) = d(gz, gx_0) \approx_{2\mathsf{D}} d(z, gx_0)$$

for all $g \in G$, where $\mathsf{D} := \mathrm{rad}(Gz) \geq \frac{1}{2}\,\mathrm{diam}(Gz)$. Therefore $zu \approx_{2\mathsf{D}} zv$ which implies $uo \approx_{2\mathsf{D}} ov$. It follows that $o \approx_{\mathsf{D}} m$, where m is the midpoint of $[u, v]$. Finally, applying Lemma 3.11 gives $p \approx_{\delta} q \approx_{3\delta} o \approx_{\mathsf{D}} m \approx_{3\delta} c$ and we are done. □

3G. ***Almost fixed point sets.*** Let G be a finite group acting by isometries on a δ-hyperbolic space $\mathcal{X}$. Given $\mathsf{R} \geq 0$, let

$$\mathrm{Fix}_{\mathcal{X}}(G, \mathsf{R}) := \{x \in \mathcal{X} \mid \mathrm{diam}(Gx) \leq \mathsf{R}\}$$

be the set of R-*almost fixed points* of G in $\mathcal{X}$.

Lemma 3.13. *The set* $\mathrm{Fix}_{\mathcal{X}}(G, 2\delta)$ *is nonempty. Moreover, if* $\mathsf{R} \geq \delta$ *then*

$$\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R}) \approx_{\mathsf{R}+\delta} \mathrm{Fix}_{\mathcal{X}}(G, 2\delta).$$

Proof. For any $x \in \mathcal{X}$, it is straightforward to check that $\mathrm{Fix}_{\mathcal{X}}(G, 2\delta)$ contains $\mathrm{circ}(Gx) \neq \varnothing$. Furthermore, if $x \in \mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R})$ then, by Lemma 3.11, we have

$$xc \leq \mathrm{rad}(Gx) \leq \tfrac{1}{2}\,\mathrm{diam}(Gx) + \delta \leq \mathsf{R} + \delta,$$

and hence $\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R}) \subseteq_{\mathsf{R}+\delta} \mathrm{Fix}_{\mathcal{X}}(G, 2\delta)$. The reverse inclusion is immediate. □

Thus, to understand the geometry of $\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R})$, for $\mathsf{R} \geq \delta$, it suffices to study that of $\mathrm{Fix}_{\mathcal{X}}(G, 2\delta)$. One can also show that $\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R})$ is quasiconvex for $\mathsf{R} \geq \delta$.

Lemma 3.14. *Let* $\mathsf{R} \geq \delta$. *For any point* $x \in \mathcal{X}$, *let* c *be a circumcenter for its* G*-orbit, and* p *be any nearest point projection to* $\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R})$. *Then* $cp \leq 2\delta + 4\mathsf{R}$.

Proof. For all $g \in G$, we have

$$d(gx, p) \leq d(gx, gp) + d(gp, p) \leq d(x, p) + 2\mathsf{R} \leq d(x, c) + 2\mathsf{R} \leq \mathrm{rad}(Gx) + 2\mathsf{R}.$$

Applying Lemma 3.10 completes the proof. □

It is worth noting that when $\mathsf{R} < \delta$, Lemmas 3.13 and 3.14 need not hold: it is possible for $\mathrm{Fix}_{\mathcal{X}}(G, 2\mathsf{R})$ to lie very deeply inside $\mathrm{Fix}_{\mathcal{X}}(G, 2\delta)$, as shown by the example below. Recall that the point $(r, \theta, t) \in [0, \infty) \times \mathbb{R} \times \mathbb{R}$ in cylindrical coordinates on $\mathbb{R}^3$ represents $(r\cos\theta, r\sin\theta, t) \in \mathbb{R}^3$ in Cartesian coordinates.

Example 3.15 (Rocketship). A *rocketship* of length $l > 0$ with $n \geq 2$ fins, denoted $\mathfrak{R} = \mathfrak{R}(n, l)$, is the union of the following sets defined using cylindrical coordinates:

- the *nose* $\mathfrak{N} = \{(t, \theta, t) \mid 0 \leq t \leq 1,\ \theta \in \mathbb{R}\}$, a right circular cone of height 1 and base radius 1;
- the *shaft* $\mathfrak{S} = \{(1, \theta, t) \mid 1 \leq t \leq l + 1,\ \theta \in \mathbb{R}\}$, a right circular cylinder of height l and base radius 1; and
- the *fins* $\mathfrak{F}_n = \{(1, 2k\pi/n, t) \mid t \geq l + 1,\ k \in \mathbb{Z}\}$, a disjoint union of n closed rays.

We endow $\mathfrak{R}$ with the induced path metric from $\mathbb{R}^3$ with the Euclidean metric. One can show that $\mathfrak{R}$ is quasi-isometric to a tree and hence δ-hyperbolic for some $\delta > 0$; this can be done by collapsing the radial component of the nose and shaft. Moreover, $\delta \geq \pi/2$ for l sufficiently large. The group $G = \mathbb{Z}/n\mathbb{Z}$ acts isometrically on $\mathfrak{R}$ by rotations about the t-axis through integral multiples of $2\pi/n$. For any $x \in \mathfrak{F}_n$, the circumcenters of Gx are points of the form $(1, (4k + 1)\pi/2n, l + 1)$, where $k \in \mathbb{Z}$. For $\mathsf{R} \geq 0$ sufficiently small, $\mathrm{Fix}_{\mathfrak{R}}(G, 2\mathsf{R})$ is contained in $\mathfrak{N}$. Therefore, $\mathrm{circ}(Gx)$ is at least a distance l away from $\mathrm{Fix}_{\mathfrak{R}}(G, 2\mathsf{R})$. Furthermore, $\mathrm{Fix}_{\mathfrak{R}}(G, 2\delta)$ contains both $\mathfrak{N}$ and $\mathfrak{S}$, and so its Hausdorff distance from $\mathrm{Fix}_{\mathfrak{R}}(G, 2\mathsf{R})$ is at least l.

4. Singular Euclidean structures

We now generalize Bowditch's [2006b] construction of singular Euclidean surfaces which are used to estimate weighted intersection numbers. Suppose $S = (S, \Omega)$ is a closed surface of genus g with a set of m marked points Ω such that $\xi(S) \geq 2$. Throughout this section, fix an n-tuple $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ of distinct multicurves in $\mathcal{C}(S)$. A vector $\mathbf{t} = (t_1, \dots, t_n) \neq \mathbf{0}$ of nonnegative real numbers shall be referred to as a *weight vector*. Write $\mathbf{t} \cdot \boldsymbol{\alpha}$ for the formal sum $\sum_i t_i \alpha_i$. For simplicity, assume that $\boldsymbol{\alpha}$ fills S and that all entries of $\mathbf{t}$ are positive. The appropriate modifications for the nonfilling case shall be dealt with in Section 7A.

4A. ***Construction of $S(\mathbf{t} \cdot \boldsymbol{\alpha})$.*** Realize the multicurves α_i on S so that they intersect generally and pairwise minimally. The union of the α_i is a connected 4-valent graph Υ on S. The closure of each component of $S - \Upsilon$ is a polygon with at most one marked point. The polygons together with Υ give S the structure of a 2-dimensional cell complex. By taking the dual 2-cell structure, we obtain a tiling of S by rectangles which are in bijection with the self-intersection points of $\boldsymbol{\alpha}$. We will insist that any marked point of S coincides with a vertex of this tiling.

Each rectangle R corresponding to an intersection of α_i with α_j is isometrically identified with a Euclidean rectangle of side lengths t_i and t_j so that α_i is transverse to the two sides of length t_i. Each vertex in this tiling meeting $k \neq 4$ corners of rectangles becomes a singular point with cone angle $k\pi/2$. This gives a singular

Euclidean metric on S. We may arrange for each α_i to be locally geodesic by requiring $\alpha_i \cap R$ to be a straight line connecting the midpoints of opposite sides of R, for every rectangle R meeting α_i. Thus, each component of α_i is the core curve of an annulus of width t_i formed by taking the union of all rectangles R it meets.

The singular Euclidean surface defined above shall be denoted $S(\mathbf{t}\cdot\boldsymbol{\alpha})$. We remark that the metric depends on the realization of $\boldsymbol{\alpha}$ on S up to isotopy, however, any such choice will work equally well for our purposes.

We will allow representatives of a curve $\gamma \in \mathcal{C}(S)$ to meet marked points to speak of (locally) geodesic representatives. Say c is a *representative* of γ if there exists an embedded curve c' representing γ and a homotopy $\mathbf{F} : S^1 \times [0,1] \to S$ such that $\mathbf{F}(\theta, 0) = c'(\theta)$, $\mathbf{F}(\theta, 1) = c(\theta)$ and $\mathbf{F}(S^1 \times \{t\}) \subseteq S - \Omega$ for all $0 \le t < 1$. A locally geodesic representative c of γ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ may not necessarily be embedded. In these cases, there is a decomposition of the circle $S^1 = \cup I_k$ into a finite union of closed intervals with disjoint interiors so that $c : S^1 \to S(\mathbf{t}\cdot\boldsymbol{\alpha})$ sends each I_k to a straight line segment with endpoints at singular points or marked points.

By a *geodesic representative* of γ, we mean a curve representing γ attaining the minimal length among all representatives of γ. Geodesic representatives exist: there is a lower bound on the injectivity radius and distance between singular points on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$, and therefore there are only finitely many locally geodesic representatives of γ with length less than any given constant. We will use $l(\gamma)$ to denote the length of a geodesic representative of γ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$.

For notational convenience, define a function on weight vectors by setting

$$\|\mathbf{t}\|_{\boldsymbol{\alpha}} := \sqrt{i(\mathbf{t}\cdot\boldsymbol{\alpha})},$$

where

$$i(\mathbf{t}\cdot\boldsymbol{\alpha}) = \sum_{j<k} t_j t_k i(\alpha_j, \alpha_k)$$

is the *self-intersection number* of $\mathbf{t}\cdot\boldsymbol{\alpha}$. This serves as a rescaling factor for the singular Euclidean surface $S(\mathbf{t}\cdot\boldsymbol{\alpha})$. We will extend intersection number linearly:

$$i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) := \sum_i t_i i(\alpha_i, \gamma).$$

Proposition 4.1. *The singular Euclidean surface $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ has the following properties*:

(1) *$S(\mathbf{t}\cdot\boldsymbol{\alpha})$ has area $\|\mathbf{t}\|_{\boldsymbol{\alpha}}^2 = \sum_{j<k} t_j t_k i(\alpha_j, \alpha_k)$.*

(2) *For all curves $\gamma \in \mathcal{C}(S)$, we have*

$$l(\gamma) \le i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) \le \sqrt{2}\, l(\gamma).$$

(3) *There exists an essential annulus on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ whose width is at least $\mathsf{W}_0\|\mathbf{t}\|_{\boldsymbol{\alpha}}$, where $\mathsf{W}_0 > 0$ is a constant depending only on $\xi(S)$.*

Here, the *width* of an annulus is the length of a shortest arc connecting its two boundary components. The first claim is immediate from the construction. The second claim shall be proven in Section 4B below; and the third in Section 7. It is worth mentioning that the third claim holds for a larger class of metrics satisfying a suitable isoperimetric inequality. The metric on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ can be approximated by a nonsingular Riemannian metric but we shall not need to do so.

4B. ***A grid structure on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$.*** A *quarter-translation surface* is a topological surface S with a finite set of singularities ς, together with an atlas of charts from $S-\varsigma$ to $\mathbb{R}^2$ whose transition maps are translations of $\mathbb{R}^2$ possibly composed with rotations through integral multiples of $\pi/2$. The singular points have cone angles which are integral multiples of $\pi/2$ and at least π.

Given a quarter-translation surface S, we may pull back the standard Euclidean metric on $\mathbb{R}^2$ to give a singular Euclidean metric on S. Geodesics which do not meet any singular points or marked points with respect to this metric can only self-intersect orthogonally. We can also define an L^1-metric on S by pulling back the metric given infinitesimally by $|dx|+|dy|$ on $\mathbb{R}^2$. We will work with the singular Euclidean metric unless otherwise specified. The following is immediate:

Lemma 4.2. *Let $l^2(\eta)$ and $l^1(\eta)$ denote, respectively, the Euclidean and L^1-lengths of a path η on S. Then $l^2(\eta)\le l^1(\eta)\le\sqrt{2}l^2(\eta)$.*

We may pull back the horizontal and vertical directions on $\mathbb{R}^2$ to give a preferred (unordered) pair of orthogonal directions on S defined away from the singular points. These shall be referred to as the *grid directions*. Geodesics which run parallel to a grid direction will be called *grid arcs*. Every nonsingular point on S has an open rectangular neighborhood, with sides parallel to the grid directions, on which the grid leaves restrict to give a pair of transverse foliations. Such a rectangle will be called an *open grid rectangle*.

It is straightforward to check that $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ is a quarter-translation surface. We will assume that the grid directions on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ run parallel to the sides of the rectangles used in its construction.

Lemma 4.3. *Given a curve $\gamma\in\mathcal{C}(S)$, let c be any of its geodesic representatives on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ with respect to the Euclidean metric. Then $l^1(c)=i(\mathbf{t}\cdot\boldsymbol{\alpha},\gamma)$.*

Proof. If c is embedded then we can isotope it to another geodesic representative meeting at least one singularity. Thus we can assume that S^1 decomposes as a finite union of intervals $\cup I_k$ with disjoint interiors such that $c: S^1\to S(\mathbf{t}\cdot\boldsymbol{\alpha})$ embeds each I_k as a straight line segment connecting singularities or marked points.

We can homotope c to a closed path $c': S^1\to S(\mathbf{t}\cdot\boldsymbol{\alpha})$ so that each $c'(I_k)$ is an edge-path in the 1-skeleton of $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ with the same endpoints as $c(I_k)$. The homotopy can be performed in a way which preserves the l^1-length of the path and without creating new intersection points with any of the α_i. One can check that c intersects each α_i minimally and thus the same is also true of c'. Finally, we deduce $l^1(c'(S^1)) = i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma)$ by observing that every edge in the 1-skeleton of $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ transverse to α_i has length t_i. □

The second claim of Proposition 4.1 follows from the previous two lemmas.

5. Hulls in the curve complex

Let $S = (S, \Omega)$ be a connected compact surface S without boundary with a finite set of marked points Ω satisfying $\xi(S) \geq 2$. Throughout this section, we will fix an n-tuple $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ of distinct multicurves in $\mathcal{C}(S)$, where $n \geq 2$. We will assume that no pair α_i and α_j has a common component. We shall establish a coarse equality between two subsets of $\mathcal{C}(S)$ determined by $\boldsymbol{\alpha}$ — its hyperbolic hull $\mathrm{Hull}(\boldsymbol{\alpha})$, defined purely in terms of the geometry of $\mathcal{C}(S)$; and $\mathrm{Short}(\boldsymbol{\alpha}, \mathsf{L})$ which is defined using only intersection numbers. We also give a combinatorial method of approximating nearest point projections to $\mathrm{Hull}(\boldsymbol{\alpha})$.

5A. ***Short curve sets.*** Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ be an n-tuple of distinct multicurves in $\mathcal{C}^0(S)$, and $\mathbf{t} = (t_1, \ldots, t_n)$ be a weight vector. Given $\mathsf{L} \geq 0$, define

$$\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L}) := \{\gamma \in \mathcal{C}(S) \mid i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) \leq \mathsf{L}\|\mathbf{t}\|_{\boldsymbol{\alpha}}\}.$$

If $\|\mathbf{t}\|_{\boldsymbol{\alpha}} = 0$ then this set is contained in the 1-neighborhood of $\boldsymbol{\alpha}$. Note that $\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ remains invariant under multiplying $\mathbf{t}$ by a positive scalar. When $\boldsymbol{\alpha}$ fills S, the geodesic length of a curve γ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ approximates its intersection number with $\mathbf{t}\cdot\boldsymbol{\alpha}$ (Proposition 4.1). (The same is also true in the nonfilling case — see Section 7A). Thus, we can view $\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ as the set of bounded length curves on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ rescaled to have unit area.

Lemma 5.1. *There exists a constant* $\mathsf{L}_0 > 0$ *depending only on* $\xi(S)$ *such that, for any* $\mathsf{L} \geq \mathsf{L}_0$, *the set* $\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ *is nonempty. Moreover*,

$$\mathrm{diam}_{\mathcal{C}(S)}(\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})) \leq 4\log_2 \mathsf{L} + \mathsf{k}_0,$$

where k_0 *is a constant depending only on* $\xi(S)$.

Consequently, up to bounded error, we can view $\mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ as a single curve in $\mathcal{C}(S)$ which has minimal intersection number with $\mathbf{t}\cdot\boldsymbol{\alpha}$. The proof of this result will be given in Section 8, and largely follows the proof of Lemma 4.1 in [Bowditch 2006b].

5B. ***A hull via intersection numbers.*** For $\mathsf{L} \geq 0$, define the L-*short curve hull* of $\boldsymbol{\alpha}$ to be

$$\mathrm{Short}(\boldsymbol{\alpha}, \mathsf{L}) := \bigcup_{\mathbf{t}} \mathrm{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L}),$$

where the union is taken over all weight vectors $\mathbf{t} \in \mathbb{R}^n_{\geq 0}$ (or, equivalently, by choosing one representative from each projective class). Write $\mathrm{Hull}(\boldsymbol{\alpha}) \subseteq \mathcal{C}(S)$ for the hyperbolic hull of $\boldsymbol{\alpha}$ considered as a set of vertices in $\mathcal{C}(S)$.

Proposition 5.2. *Let $\mathsf{L} \geq \mathsf{L}_0$. Then for any n-tuple of multicurves $\boldsymbol{\alpha}$ in $\mathcal{C}(S)$,*

$$\mathrm{Short}(\boldsymbol{\alpha}, \mathsf{L}) \approx_{\mathsf{k}_1} \mathrm{Hull}(\boldsymbol{\alpha}),$$

where k_1 depends only on $\xi(S)$, n and L.

This is essentially an extension of Bowditch's coarse description of geodesics in $\mathcal{C}(S)$ using intersection numbers, which we now reformulate:

Lemma 5.3 [Bowditch 2006b, Proposition 6.2]. *Let $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2)$ be a pair of multicurves in $\mathcal{C}(S)$. Let $[\alpha_1, \alpha_2]$ denote any geodesic segment connecting α_1 and α_2 in $\mathcal{C}(S)$. Then for all $\mathsf{L} \geq \mathsf{L}_0$, we have*

$$\mathrm{Short}(\boldsymbol{\alpha}', \mathsf{L}) \approx_{\mathsf{k}'_1} [\alpha_1, \alpha_2],$$

where $\mathsf{k}'_1 \geq 0$ depends only $\xi(S)$ and L.

Proof of Proposition 5.2. Applying Lemma 5.3 to all pairs of multicurves (α_i, α_j) in $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, we obtain the inclusion:

$$\mathrm{Hull}(\boldsymbol{\alpha}) \subseteq_{\mathsf{k}'_1} \mathrm{Short}(\boldsymbol{\alpha}, \mathsf{L}).$$

Let $\mathbf{t} = (t_1, \ldots, t_n)$ be a weight vector and assume, without loss of generality, that the quantity $t_j t_k i(\alpha_j, \alpha_k)$ is maximized when $\{j, k\} = \{1, 2\}$. Let $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2)$ and $\mathbf{t}' = (t_1, t_2)$. Since there are $n(n-1)/2$ distinct unordered pairs of indices $\{j, k\}$, it follows that

$$\|\mathbf{t}\|^2_{\boldsymbol{\alpha}} = \sum_{j<k} t_j t_k i(\alpha_j, \alpha_k) \leq \frac{n(n-1)}{2} t_1 t_2 i(\alpha_1, \alpha_2) = \frac{n(n-1)}{2} \|\mathbf{t}'\|^2_{\boldsymbol{\alpha}'}.$$

Now let γ be a curve in $\mathrm{short}(\mathbf{t} \cdot \boldsymbol{\alpha}, \mathsf{L})$. Then

$$i(\mathbf{t}' \cdot \boldsymbol{\alpha}', \gamma) \leq i(\mathbf{t} \cdot \boldsymbol{\alpha}, \gamma) \leq \mathsf{L}\|\mathbf{t}\|_{\boldsymbol{\alpha}} \leq \mathsf{L}\sqrt{\frac{n(n-1)}{2} \|\mathbf{t}'\|^2_{\boldsymbol{\alpha}'}} \leq \frac{n\mathsf{L}}{\sqrt{2}} \|\mathbf{t}'\|_{\boldsymbol{\alpha}'},$$

which implies

$$\mathrm{short}(\mathbf{t} \cdot \boldsymbol{\alpha}, \mathsf{L}) \subseteq \mathrm{short}\Big(\mathbf{t}' \cdot \boldsymbol{\alpha}', \frac{n\mathsf{L}}{\sqrt{2}}\Big).$$

Invoking Lemma 5.3, we have

$$\mathrm{short}\Big(\mathbf{t}' \cdot \boldsymbol{\alpha}', \frac{n\mathsf{L}}{\sqrt{2}}\Big) \subseteq_{\mathsf{r}} [\alpha_1, \alpha_2] \subseteq \mathrm{Hull}(\boldsymbol{\alpha}),$$

where $\mathsf{r} \geq 0$ is some constant depending on n, L and $\xi(S)$. $\square$

We can describe the above proof in terms of the geometry of $S(\mathbf{t}\cdot\boldsymbol{\alpha})$. Assume $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ has unit area. One can obtain $S(\mathbf{t}'\cdot\boldsymbol{\alpha}')$ by homotoping the annuli consisting of rectangles traversed by α_i to the core curve α_i for each $i \neq 1, 2$. The maximality assumption on α_1 and α_2 ensures that the total area of the remaining rectangles is at least $2/(n(n-1))$. Scale $S(\mathbf{t}'\cdot\boldsymbol{\alpha}')$ by a factor of at most $n/\sqrt{2}$ to give it unit area. This process scales the length of a curve γ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ by a factor of at most $n/\sqrt{2}$.

5C. ***Nearest point projections to hulls.*** In this section, we approximate nearest point projections to short curve hulls using only intersection number conditions.

Definition 5.4. Let $\beta \in \mathcal{C}(S)$ be a multicurve. A weight vector $\mathbf{t} = (t_1, \dots, t_n)$ satisfying

$$t_j i(\alpha_j, \beta) = t_k i(\alpha_k, \beta)$$

for all j, k is called a *balance vector* for β with respect to $\boldsymbol{\alpha}$.

If β intersects all α_i then setting $t_i = i(\alpha_i, \beta)^{-1}$ yields the unique balance vector up to positive scale. If not, we can set $t_i = 1$ whenever $i(\alpha_i, \beta) = 0$ and $t_i = 0$ otherwise to produce a balance vector. Let $\mathbf{t}_\beta$ denote any balance vector for β. We also remark that the above definition is analogous to the notion of *balance time* for quadratic differentials as described by Masur and Minsky [1999].

The proof of the following will be given at the end of this section.

Proposition 5.5. *Assume* $\mathsf{L} \geq \mathsf{L}_0$. *Given a multicurve* $\beta \in \mathcal{C}(S)$, *let* γ *be any nearest point projection of* β *to* $\mathrm{Hull}(\boldsymbol{\alpha})$. *Then*

$$\gamma \approx_{\mathsf{k}_2} \mathrm{short}(\mathbf{t}_\beta \cdot \boldsymbol{\alpha}, \mathsf{L}),$$

where $\mathsf{k}_2 \geq 0$ *depends only on* $\xi(S)$, n *and* L.

As was the case with Proposition 5.2, this is an extension of a result of Bowditch. His result was originally phrased in terms of centers for geodesic triangles; however, our statement agrees with it up to uniformly bounded error.

Lemma 5.6 [Bowditch 2006b, Proposition 3.1 and Section 4]. *Let* α_1, α_2 *and* β *be multicurves in* $\mathcal{C}(S)$. *Let* $\mathbf{t}'_\beta$ *be a balance vector for* β *with respect to* $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2)$. *Let* γ *be a nearest point projection of* β *to* $[\alpha_1, \alpha_2]$. *Then*

$$\gamma \approx_{\mathsf{k}'_2} \mathrm{short}(\mathbf{t}'_\beta \cdot \boldsymbol{\alpha}', \mathsf{L}),$$

where k'_2 *depends only on* $\xi(S)$ *and* L.

If β is disjoint from some α_i then Proposition 5.5 follows immediately from Lemma 2.2. We will henceforth assume this is not the case. We reduce the problem of finding a nearest point projection to a hyperbolic hull to that of projecting to a suitable geodesic.

Lemma 5.7. *Let U be a subset of a δ-hyperbolic space $\mathcal{X}$. Fix a point $w \in \mathcal{X}$. Assume there exist $x, y \in U$ and $\mathsf{R} \geq 0$ such that*

$$d_{\mathcal{X}}([x, y], [z, w]) \leq \mathsf{R}$$

for all $z \in U$. Let p and q be nearest point projections of w to $\mathrm{Hull}(U)$ *and $[x, y]$ respectively. Then*

$$p \approx_{\mathsf{R}'} q,$$

where R' depends only on R and δ.

Proof. By Lemma 3.5, it suffices to show that for all $u \in \mathrm{Hull}(U)$, any geodesic $[w, u]$ must pass within a bounded distance of q. If u lies on a geodesic segment $[z, z']$ for some $z, z' \in U$ then $[w, u]$ must lie inside the 2δ-neighborhood of $[w, z]$ or $[w, z']$. Hence, we only need to bound $d(q, [w, z])$ for all $z \in U$ in terms of δ and R. Recall that geodesic segments are δ-quasiconvex. Choose points $v \in [x, y]$ and $v' \in [z, w]$ so that $vv' = d_{\mathcal{X}}([x, y], [z, w]) \leq \mathsf{R}$. Then

$$q \subseteq_{3\delta} [w, v] \subseteq_{\mathsf{R}+\delta} [w, v'] \subseteq [w, z],$$

where we have applied Lemma 3.4 for the first comparison. □

In order to exploit the above result, we recall yet another lemma of Bowditch:

Lemma 5.8 [Bowditch 2006b, Proposition 6.3]. *Suppose $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathcal{C}(S)$ are multicurves which satisfy*

$$i(\alpha_1, \alpha_4) i(\alpha_2, \alpha_3) \leq \mathsf{r} i(\alpha_1, \alpha_2) i(\alpha_3, \alpha_4)$$

for some $\mathsf{r} > 0$. Then

$$d_S([\alpha_1, \alpha_2], [\alpha_3, \alpha_4]) \leq \mathsf{R},$$

where $\mathsf{R} \geq 0$ depends only on r and $\xi(S)$.

Proof of Proposition 5.5. Let $\mathbf{t}_\beta$ be a balance vector for β with respect to $\boldsymbol{\alpha}$. To simplify notation, assume $t_j t_k i(\alpha_j, \alpha_k)$ is maximized when $\{j, k\} = \{1, 2\}$. Then

$$t_2 t_j i(\alpha_2, \alpha_j) \leq t_1 t_2 i(\alpha_1, \alpha_2)$$

for any $j = 1, \ldots, n$. As β is assumed to intersect all the α_i, we have $t_i = i(\alpha_i, \beta)^{-1}$ (after rescaling) and so

$$i(\alpha_1, \beta) i(\alpha_2, \alpha_j) \leq i(\alpha_1, \alpha_2) i(\alpha_j, \beta).$$

Invoking Lemma 5.8 gives

$$d_S([\alpha_1, \alpha_2], [\alpha_j, \beta]) \leq \mathsf{R}.$$

Let γ_{12} and γ be nearest point projections of β to $[\alpha_1, \alpha_2]$ and Hull($\boldsymbol{\alpha}$) respectively. Applying Lemma 5.7 with $U = \boldsymbol{\alpha},\ x = \alpha_1,\ y = \alpha_2$, and $w = \beta$ gives

$$d_S(\gamma_{12}, \gamma) \le \mathsf{R}',$$

where R' depends only on $\xi(S)$.

Now suppose γ' is a curve in short($\mathbf{t}_\beta \cdot \boldsymbol{\alpha}$, L). Using the same reasoning as for the proof of Proposition 5.2, we see that

$$\gamma' \in \mathrm{short}(\mathbf{t}_\beta \cdot \boldsymbol{\alpha}, \mathsf{L}) \subseteq \mathrm{short}\Big(\mathbf{t}'_\beta \cdot \boldsymbol{\alpha}', \frac{n\mathsf{L}}{\sqrt{2}}\Big),$$

where $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2)$ and $\mathbf{t}'_\beta = (t_1, t_2)$. By Lemma 5.6, we deduce that

$$d_S(\gamma', \gamma_{12}) \le \mathsf{k}'_2$$

for some k'_2 depending only on n, L and $\xi(S)$. The preceding inequalities give

$$d_S(\gamma', \gamma) \le \mathsf{R}' + \mathsf{k}'_2,$$

which concludes the proof of the proposition. □

6. Covering maps

6A. ***Operations on curves arising from covering maps.*** We first recall some definitions and notation. Let $P : \Sigma \to S$ be a finite degree covering map of surfaces. The preimage $P^{-1}(a)$ of a simple closed curve a on S under P is a multicurve on Σ. This induces a one-to-many *lifting map* $\Pi : \mathcal{C}(S) \to \mathcal{C}(\Sigma)$ between curve complexes by setting $\Pi(a) := P^{-1}(a)$. Recall the following theorem of Rafi and Schleimer:

Theorem 1.1 [Rafi and Schleimer 2009]. *Let $P : \Sigma \to S$ be a finite degree covering map. Then the map $\Pi : \mathcal{C}(S) \to \mathcal{C}(\Sigma)$ defined above is a Λ-quasi-isometric embedding*, *where Λ depends only on $\xi(\Sigma)$ and* deg *P.*

It immediately follows that $\Pi(\mathcal{C}(S))$ is quasiconvex in $\mathcal{C}(\Sigma)$. This naturally leads to the question of understanding nearest point projections to $\Pi(\mathcal{C}(S))$. Define an operation $\pi : \mathcal{C}(\Sigma) \to \Pi(\mathcal{C}(S))$ as follows: given a curve $\alpha \in \mathcal{C}(\Sigma)$, let $b \in \mathcal{C}(S)$ be a curve which has minimal intersection number with $P(\alpha)$ on S and set $\pi(\alpha) = \Pi(b)$.

Theorem 6.1. *Let $P : \Sigma \to S$ be a finite degree covering map*, *and let Π and π be as above. Given a curve $\alpha \in \mathcal{C}(\Sigma)$*, *let γ be a nearest point projection of α to $\Pi(\mathcal{C}(S))$ in $\mathcal{C}(\Sigma)$. Then $\pi(\alpha) \approx_{\mathsf{k}_3} \gamma$*, *where k_3 depends only on* deg *P and $\xi(\Sigma)$.*

Consequently, the operation $\alpha \mapsto \pi(\alpha)$ is coarsely well defined. The above will be proven in Section 6B, and the following in Section 6C.

Proposition 6.2. *Suppose further that P is regular, and let G be its group of deck transformations. Let γ' be a circumcenter of the G-orbit of a curve α in $\mathcal{C}(\Sigma)$. Then $\pi(\alpha) \approx_{\mathsf{k}_4} \gamma'$, where k_4 is some constant depending only on* $\deg P$ *and* $\xi(\Sigma)$.

Recall that the *deck transformation group* $\mathrm{Deck}(P)$ of a covering map $P : \Sigma \to S$ is the group of all homeomorphisms $f \in \mathrm{Homeo}(\Sigma)$ satisfying $P \circ f = P$. In order for the above statement to make sense, we must check that $\mathrm{Deck}(P)$ can be identified with its image in the *mapping class group* $\mathrm{Mod}(\Sigma) = \mathrm{Homeo}(\Sigma)/\mathrm{Homeo}_0(\Sigma)$.

Lemma 6.3. *Suppose S has negative Euler characteristic, and let $P : \Sigma \to S$ be a finite degree covering map. Then the natural map* $\mathrm{Deck}(P) \to \mathrm{Mod}(\Sigma)$ *is injective.*

Proof. We will only give a sketch proof. Endow $\mathrm{int}(S)$ with a hyperbolic metric and pull it back to $\mathrm{int}(\Sigma)$ via P. The group $\mathrm{Deck}(P)$ then acts on $\mathrm{int}(\Sigma)$ by isometries. The result follows since any isometry of a hyperbolic surface isotopic to the identity must in fact coincide with the identity. □

Note, however, that this lemma does not hold for covers of the torus or annulus.

6B. *Nearest point projections.*

6B.1. *Regular covers.* We shall first deal with the case where $P : \Sigma \to S$ is a regular cover. Let $G = \mathrm{Deck}(P)$. Given a curve $\alpha \in \mathcal{C}(\Sigma)$, observe that the set of lifts of $P(\alpha)$ to Σ via P is exactly $G\alpha$. Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ be an n-tuple of curves whose entries are the lifts of $P(\alpha)$ in any order. Note that $n \geq 1$ is some divisor of $\deg P$. Let $\mathbf{1}$ denote the vector of length n with all entries equal to 1.

Lemma 6.4. *Let α and $\boldsymbol{\alpha}$ be as above. Then $\pi(\alpha) \in$* $\mathrm{short}(\mathbf{1} \cdot \boldsymbol{\alpha}, \mathsf{L}_0|G|)$ *where L_0 is a constant depending only on $\xi(\Sigma)$.*

Proof. Let b be a closed curve on S. Each point of $b \cap P(\alpha)$ on S lifts to exactly $|G| = \deg P$ points of $P^{-1}(b) \cap G\alpha$ on Σ via P, hence

$$i(P^{-1}(b), \boldsymbol{\alpha}) = |G| i(b, P(\alpha)).$$

By Lemma 5.1, there exists a curve $\gamma \in \mathcal{C}(\Sigma)$ such that

$$i(\gamma, \alpha) \leq \mathsf{L}_0 \|\mathbf{1}\|_{\boldsymbol{\alpha}}$$

for some constant $\mathsf{L}_0 = \mathsf{L}_0(\xi(\Sigma))$. Now assume b has minimal intersection with $P(\alpha)$ out of all curves on S. It follows that

$$i(P^{-1}(b), \boldsymbol{\alpha}) = |G| i(b, P(\alpha)) \leq |G| i(P(\gamma), P(\alpha)) = i(G\gamma, \boldsymbol{\alpha}) \leq |G| i(\gamma, \boldsymbol{\alpha}).$$

Finally, by combining the preceding inequalities, we see that

$$i(\pi(\alpha), \boldsymbol{\alpha}) = i(P^{-1}(b), \boldsymbol{\alpha}) \leq |G| i(\gamma, \boldsymbol{\alpha}) \leq |G| \mathsf{L}_0 \|\mathbf{1}\|_{\boldsymbol{\alpha}}.$$

Thus $\pi(\alpha) \in \mathrm{short}(\mathbf{1} \cdot \boldsymbol{\alpha}, \mathsf{L})$ for $\mathsf{L} = \mathsf{L}_0|G|$. □

Lemma 6.5. *Given $\gamma \in \Pi(\mathcal{C}(S))$, let β be any of its nearest point projections to* $\mathrm{Hull}(\boldsymbol{\alpha})$. *Then $d_\Sigma(\pi(\alpha), \beta) \leq \mathsf{k}_5$, where k_5 depends only on* $\deg P$ *and* $\xi(\Sigma)$.

Proof. We may replace γ with the multicurve $G\gamma$ since their nearest point projections to $\mathrm{Hull}(\boldsymbol{\alpha})$ are a uniformly bounded distance apart. Since G acts transitively on $G\alpha$, it follows that $i(G\gamma, \alpha_i) = i(G\gamma, \alpha_j)$ for all i, j. Thus, $\mathbf{1}$ serves as a balance vector for $G\gamma$ with respect to $\boldsymbol{\alpha}$. By Proposition 5.5, we deduce that

$$\beta \approx_{\mathsf{k}_2} \mathrm{short}(\mathbf{1} \cdot \boldsymbol{\alpha}, \mathsf{L}),$$

where k_2 depends only on $\xi(\Sigma)$, n and $\mathsf{L} \geq \mathsf{L}_0$. Applying the previous lemma completes the proof. □

Proof of Theorem 6.1 for regular covers. Let α and $\boldsymbol{\alpha}$ be as above. Let γ be any curve in $\Pi(\mathcal{C}(S))$. Since $\mathrm{Hull}(\boldsymbol{\alpha})$ is quasiconvex, Lemmas 3.4 and 6.5 imply that any geodesic connecting α to γ in $\mathcal{C}(\Sigma)$ must pass within a distance r of $\pi(\alpha)$, where r depends only on $\deg P$ and $\xi(\Sigma)$. Therefore $\pi(\alpha)$ is an r-entry point of α to $\Pi(\mathcal{C}(S))$. Since $\Pi(\mathcal{C}(S))$ is also quasiconvex, Lemma 3.5 implies $\pi(\alpha)$ is a uniformly bounded distance away from any nearest point projection of α to $\Pi(\mathcal{C}(S))$. □

6B.2. *The general case.* The main obstacle in proving Theorem 6.1 for a nonregular cover $P : \Sigma \to S$ is the following: given a simple closed curve $\alpha \in \mathcal{C}(\Sigma)$ there may be some lifts of $P(\alpha)$ to Σ which are not simple. To address this issue, we pass to a suitable finite cover of Σ using a standard group theoretic argument.

Lemma 6.6. *Let $P : \Sigma \to S$ be a covering map of finite degree. Then there exists a cover $Q : \widehat{\Sigma} \to \Sigma$ such that $F := P \circ Q$ is regular and* $\deg F \leq (\deg P)!$.

Proof. Let H be the finite index subgroup of $\Gamma = \pi_1(S)$ corresponding to the covering map P, and let H_0 be the intersection of all Γ-conjugates of H. It is straightforward to check that H_0 is exactly the kernel of the action of Γ on the set of left cosets of H by left multiplication. The desired result then follows. □

The covering map F defined above is universal in the sense that any regular cover of S which factors through P must also factor through F.

Lemma 6.7. *Let $P : \Sigma \to S$ and $F : \widehat{\Sigma} \to S$ be as above. If α is a simple closed curve on Σ then all lifts of $P(\alpha)$ to $\widehat{\Sigma}$ via F are simple.*

Proof. Any lift of α to $\widehat{\Sigma}$ via Q is also a simple lift of $P(\alpha)$ via F. Since F is regular, it follows that all other lifts of $P(\alpha)$ to $\widehat{\Sigma}$ are simple. □

Let $\Phi : \mathcal{C}(S) \to \mathcal{C}(\widehat{\Sigma})$ and $\Psi : \mathcal{C}(\Sigma) \to \mathcal{C}(\widehat{\Sigma})$ be the lifting maps induced by the covering maps F and Q respectively. Let $\phi : \mathcal{C}(\widehat{\Sigma}) \to \Phi(\mathcal{C}(S))$ be the projection map associated to F as described in Section 6A. We may assume $\phi \circ \Psi = \Psi \circ \pi$.

Proof of Theorem 6.1. Given $\alpha \in \mathcal{C}(\Sigma)$, let $\hat{\alpha}$ be any of its lifts to $\widehat{\Sigma}$ via Q. Note that $\phi(\hat{\alpha}) = \Psi(\pi(\alpha))$. Let $\hat{\gamma}$ be a nearest point projection of $\hat{\alpha}$ to $\Phi(\mathcal{C}(S))$ in $\mathcal{C}(\widehat{\Sigma})$ and let $\gamma = Q(\hat{\gamma}) \in \Pi(\mathcal{C}(S))$. Since F is regular, we can apply Theorem 6.1 for regular covers to deduce that

$$d_{\widehat{\Sigma}}(\phi(\hat{\alpha}), \hat{\gamma}) \leq \hat{\mathsf{k}}_3,$$

where $\hat{\mathsf{k}}_3$ depends only on $\deg F$ and $\xi(\widehat{\Sigma})$ which can in turn be bounded in terms of $\deg P$ and $\xi(\Sigma)$. By Theorem 1.1, Ψ is a Λ-quasi-isometric embedding, where $\Lambda = \Lambda(\deg F, \xi(\widehat{\Sigma}))$, and so

$$d_{\Sigma}(\pi(\alpha), \gamma) \leq \Lambda\hat{\mathsf{k}}_3 + \Lambda.$$

By Lemma 3.6, γ is a uniformly bounded distance away from any nearest point projection of α to $\Pi(\mathcal{C}(S))$ in $\mathcal{C}(\Sigma)$ and we are done. □

6C. ***Circumcenters for regular covers.*** Let $P : \Sigma \to S$ be a regular cover, and G its deck group. Given $\alpha \in \mathcal{C}(\Sigma)$, we show $\pi(\alpha)$ approximates $\operatorname{circ}(G\alpha)$ in $\mathcal{C}(\Sigma)$.

Proof of Proposition 6.2. Since $\pi(\alpha)$ is a G-invariant multicurve, we deduce that $\operatorname{rad}(G\pi(\alpha)) \leq 1$. Proposition 5.2 and Lemma 6.4 together give

$$d_{\Sigma}(\pi(\alpha), \operatorname{Hull}(G\alpha)) \leq \mathsf{k}'_4,$$

where k'_4 depends only on $\xi(\Sigma)$ and $\deg P$. Finally, combining the above with Lemma 3.12 yields $d_{\Sigma}(\pi(\alpha), \operatorname{circ}(G\alpha)) \leq \mathsf{k}'_4 + 7\delta + 1$ as desired. □

Observe that the vertices in $\operatorname{Fix}_{\mathcal{C}(\Sigma)}(G, 1)$ coincide exactly with those of $\Pi(\mathcal{C}(S))$. An immediate corollary of Theorem 6.1 and Proposition 6.2 is the following:

Corollary 6.8. *Any circumcenter for the G-orbit of a curve $\alpha \in \mathcal{C}(S)$ is within a uniformly bounded distance of any nearest point projection of α to $\Pi(\mathcal{C}(S))$.*

Therefore Lemma 3.14 still holds for $\operatorname{Fix}_{\mathcal{C}(\Sigma)}(G, 1)$, albeit with weaker control over the constants. As Example 3.15 demonstrates, this cannot be proven using purely synthetic methods assuming only δ-hyperbolicity of $\mathcal{C}(\Sigma)$. In conclusion: "There are no rocketships in the curve complex."

7. An isoperimetric inequality on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$

7A. ***Constructing $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ for nonfilling curves.*** We now generalize the construction of $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ to encompass nonfilling curves. Assume $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ is an n-tuple of distinct multicurves and $\mathbf{t} = (t_1, \ldots, t_n) \neq \mathbf{0}$ is a weight vector satisfying $\|\mathbf{t}\|_{\boldsymbol{\alpha}} \neq 0$. Realize $\boldsymbol{\alpha}$ minimally on S to form a 4-valent graph Υ on S.

Let $\Sigma \subseteq S$ be the (possibly disconnected) subsurface filled by Υ. This can be obtained by taking a closed regular neighborhood of Υ on S and then attaching all complementary regions which are discs with at most one marked point. If $\boldsymbol{\alpha}$ fills S then $\Sigma = S$. In general, Σ will be a disjoint union of surfaces $\Sigma_1 \cup \ldots \cup \Sigma_s$. Observe that $s \le \xi(S)$ since we can find a multicurve on S so that exactly one component is contained in each Σ_k (by choosing a suitable subset of all curves appearing in $\boldsymbol{\alpha}$, for example). Some of these components may be annuli — this occurs precisely when a multicurve α_i has a component disjoint from all other α_j. All other components will have genus at least one, or are spheres where the sum of the number of marked points and boundary components is at least four.

We now define a 2-dimensional complex $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ as a quotient of S. Suppose Σ_k is an annular component of Σ whose core curve is a component of α_i. Identify Σ_k with $S^1 \times [0, t_i]$, then collapse the first coordinate to give a closed interval I_k of length t_i. Next, collapse every complementary component of Σ in S to a marked point. These marked points will be called *essential*. We then apply the construction from Section 4A to the image of each nonannular component of Σ in the quotient space. The resulting space is a finite collection of singular Euclidean surfaces and closed intervals identified along appropriate essential marked points. Note that this construction agrees with the one given in Section 4A for the case of filling curves. For brevity, call the image of a component of Σ a *component* of $S(\mathbf{t}\cdot\boldsymbol{\alpha})$.

Let c be a representative of a curve $\gamma \in \mathcal{C}(S)$ on S. Its image $\bar{c}$ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ will be a closed curve or a union of paths connecting essential marked points. Define $l(\gamma)$ to be the minimal length of $\bar{c}$ over all representatives c of γ.

Proposition 7.1. *Suppose $\boldsymbol{\alpha}$ and $\mathbf{t}$ satisfy $\|\mathbf{t}\|_{\boldsymbol{\alpha}} > 0$. Then the first two claims of Proposition 4.1 hold for $S(\mathbf{t}\cdot\boldsymbol{\alpha})$.*

The proof of the above proceeds in the same manner as for the case of filling curves. It remains to prove an analogue of the third claim.

7B. ***An isoperimetric inequality.*** Let $S = (S, \Omega)$ be a closed singular Riemannian surface S with a finite set of marked points Ω. Let Δ be a closed disc and suppose $\iota : \Delta \to S$ is a piecewise smooth immersion which restricts to an embedding on its interior. Let D denote the image $\iota(\mathrm{int}(\Delta))$.

Definition 7.2. An open disc D arising in the above manner is called a *trivial region* on S if it contains at most one marked point. The boundary ∂D is an embedded Eulerian graph on S whose edges are piecewise smooth arcs. Define $\mathrm{length}(\partial D)$ to be the sum of the lengths of these arcs using the metric on S.

Bowditch defines trivial regions as open discs on S containing at most one marked point without any conditions concerning piecewise smooth embeddings. Nevertheless, his proof of the following proposition still holds with our definition:

Proposition 7.3 [Bowditch 2006b]. *Suppose $f : [0, \infty) \to [0, \infty)$ is a homeomorphism. Let ρ be a singular Riemannian metric on an orientable closed surface S*

with unit area. Let Ω be a finite set of marked points on S. We will assume $|\Omega| \geq 5$ whenever S is a 2-sphere. If $\operatorname{area}(D) \leq f(\operatorname{length}(\partial D))$ *for any trivial region D then there is an essential annulus $A \subseteq S - \Omega$ such that* $\operatorname{width}(A) \geq \mathsf{W}_0$, *where* $\mathsf{W}_0 > 0$ *depends only on $\xi(S)$ and f.*

This section will be devoted to proving the following lemma which, together with the above proposition, implies the third claim of Proposition 4.1:

Lemma 7.4. *Suppose D is a trivial region on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$. Then*

$$\operatorname{area}(D) \leq 4 \operatorname{length}(\partial D)^2.$$

Before launching into the details of the proof, we briefly outline our argument. First, we reduce the problem to that of studying embedded closed discs on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ whose boundary is a finite union of grid arcs. We then show that such a disc D can be tiled by grid rectangles. This tiling is dual to a collection of arcs on D, where each arc is parallel to a component of some $\alpha_i \cap D$. We call the union of all rectangles meeting a given arc a *band*. The key step is to observe that any two arcs in the collection intersect at most twice. Thus, the intersection of two distinct bands is the union of at most two rectangles arising from the tiling. Conversely, any rectangle from the tiling is contained in the intersection of two such bands. We can then bound the area of the rectangles in terms of the length of ∂D.

7B.1. *Technical adjustments.* Let us first make a couple of observations to simplify the problem.

Lemma 7.5. *Any trivial region D on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ can be perturbed to a trivial region D' whose boundary is a finite union of grid leaves. Moreover, D' can be chosen so that* $\operatorname{area}(D') \geq \operatorname{area}(D)$ *and* $\operatorname{length}(\partial D') \leq \sqrt{2} \operatorname{length}(\partial D)$.

We will henceforth assume that the boundary of any trivial region on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ is a finite union of grid leaves.

Let $\iota : \Delta \to S(\mathbf{t} \cdot \boldsymbol{\alpha})$ be a piecewise smooth immersion whose restriction to $\operatorname{int}(\Delta)$ is an embedding with image D. Observe that $\iota : \partial \Delta = S^1 \to \partial D$ is an immersion of a circle which runs over each edge of ∂D at most twice. We will metrize Δ by pulling back the metric on $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ via ι.

Lemma 7.6. *Suppose D and Δ are as given above. Then* $\operatorname{area}(\Delta) = \operatorname{area}(D)$ *and* $\operatorname{length}(\partial D) \leq \operatorname{length}(\partial \Delta) \leq 2 \operatorname{length}(\partial D)$.

7B.2. *Tiling Δ by rectangles.* The disc Δ inherits grid directions from $S(\mathbf{t} \cdot \boldsymbol{\alpha})$ via ι away from the preimage of the singular points. The boundary decomposes as a finite union $\partial \Delta = \cup I_k$ of closed grid arcs with disjoint interiors. We may assume that this decomposition is minimal, that is, it cannot be obtained from any other such decomposition by subdividing arcs. An endpoint of any grid arc I_k will be called a *corner point* of $\partial\Delta$. A corner point which does not coincide with a singularity or a marked point must be an orthogonal intersection point of two grid arcs.

It is worth noting that $\partial\Delta$ must contain at least two corner points and at least three if D contains no marked points. To see this, recall that the grid leaves on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ are parallel to some α_i. Any of the forbidden cases will imply that some α_i is trivial, peripheral, self-intersects or does not intersect some α_j minimally.

Let us refer to marked points, corner points and singularities collectively as *bad points*. Let $Z \subset \Delta$ be the union of $\partial\Delta$ with all grid arcs in Δ which have a bad point for at least one of their endpoints. Since there are finitely many bad points in Δ, it follows that Z is a finite embedded graph on Δ. A vertex $v \in \operatorname{int}\Delta \cap Z$ has valence k if and only if the cone angle at v is $k\pi/2$. If v is a vertex which lies on $\partial\Delta$ then it has valence $k+1$ if and only if the cone angle at v inside Δ is $k\pi/2$. It follows that every vertex v of Z has valence at least 2, and at least 3 if v is not a marked point.

Lemma 7.7. *There exists a tiling of Δ by finitely many grid rectangles with Z as its* 1*-skeleton.*

Proof. First note that there are finitely many connected components of $\Delta - Z$ since Z is a finite graph. Let R be such a component and let $\bar{R}$ be its completion with respect to its induced path metric. Observe that $\bar{R}$ is a closed planar region admitting a Euclidean metric with piecewise geodesic boundary, where the interior angle between adjacent edges of $\partial\bar{R}$ is $\pi/2$. By the Gauss–Bonnet formula, the sum of its interior angles must equal $2\pi\chi(R)$. Since the frontier of R in Δ meets at least one vertex of Z, the angle sum must be strictly positive. As R is planar, it follows that $\chi(R) = 1$ and therefore $\bar{R}$ is a Euclidean rectangle. Also note that Z is connected, for otherwise there would exist some component of $\Delta - Z$ with disconnected frontier.

The inclusion $R \hookrightarrow \Delta$ can be extended continuously to a map $\bar{R} \to \Delta$, sending each edge of $\partial\bar{R}$ isometrically to an edge of Z meeting the frontier of R. Thus R is a grid rectangle since the edges of Z, by construction, are parallel to the grid directions. Finally, the closures of distinct rectangles R and R' can only intersect in a union of vertices and edges of Z. $\square$

7B.3. *Controlling the area.* Let $\mathcal{A}$ be the set of maximal grid arcs in Δ which intersect Z only at midpoints of edges of Z. This is a collection of arcs dual to the rectangular tiling of Δ as described in Lemma 7.7. For any arc $a \in \mathcal{A}$, there is curve $\alpha \in \boldsymbol{\alpha}$ such that a can be properly isotoped in Δ to a component of $\iota^{-1}(\alpha \cap D)$ without passing through any singular points or marked points. (There cannot be any closed curves in Δ dual to the tiling as this would imply some α_i is not essential.) Let $B = B(a)$ be the union of all rectangles in the tiling which meet a. We will call B a *band* and a a *core arc* of B. Define $\mathrm{width}(B)$ to be the length of any edge of Z crossed by a. The set of bands in Δ is in bijection with $\mathcal{A}$.

Lemma 7.8. *The intersection of two distinct bands B and B' is the union of at most two rectangles whose side lengths are* $\mathrm{width}(B)$ *by* $\mathrm{width}(B')$. *Conversely, each rectangle in the tiling lies in the intersection of a unique pair of distinct bands.*

Proof. Let a and a' be core arcs of B and B' respectively. If a and a' intersect at least 3 times then they must bound a bigon in Δ containing no marked points. We can properly isotope a and a' in Δ to components of $\iota^{-1}(\alpha_i \cap D)$ and $\iota^{-1}(\alpha_j \cap D)$, for some α_i and α_j respectively, without passing through any singular points or marked points. Since any right-angled bigon on Δ must contain at least one singularity, it follows that α_i and α_j also bound a bigon in D, contradicting minimality.

For the converse, simply take the bands corresponding to the unique pair of arcs which have an intersection point inside the given rectangle. □

We will refer to an edge of Z lying in $\partial\Delta$ simply as an *edge* of $\partial\Delta$.

Lemma 7.9. *Let Δ be as above. Then*

$$\mathrm{area}(\Delta) \leq \tfrac{1}{2}\,\mathrm{length}(\partial\Delta)^2.$$

Proof. By Lemma 7.8, Δ is a union of rectangles, each of which lies in the intersection of a pair of distinct bands. Thus

$$\mathrm{area}(\Delta) = \mathrm{area}\Bigl(\bigcup_{B\neq B'} B \cap B'\Bigr) = \sum_{B\neq B'} \mathrm{area}(B\cap B').$$

Since the intersection of two distinct bands is the union of at most two rectangles whose side lengths are equal to the widths of the bands, we have

$$\mathrm{area}(B\cap B') \leq 2\,\mathrm{width}(B)\times\mathrm{width}(B'),$$

and hence

$$\mathrm{area}(\Delta) \leq 2\sum_{B\neq B'} \mathrm{width}(B)\times\mathrm{width}(B') \leq 2\Bigl(\sum_{B}\mathrm{width}(B)\Bigr)^2.$$

Finally, the desired result follows from observing that

$$\mathrm{length}(\partial\Delta) = 2\sum_{B}\mathrm{width}(B),$$

where the sum is taken over all bands B in Δ. □

Combining this with Lemmas 7.5 and 7.6 completes the proof of Lemma 7.4.

8. Proof of Lemma 5.1

Fix an n-tuple of distinct multicurves $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ and a weight vector $\mathbf{t} = (t_1, \dots, t_n) \neq \mathbf{0}$. We show that for all $\mathsf{L} \geq \mathsf{L}_0$, where L_0 is to be determined,

$$\operatorname{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L}) = \{\gamma \in \mathcal{C}(S) \mid i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) \leq \mathsf{L}\|\mathbf{t}\|_{\boldsymbol{\alpha}}\}$$

is nonempty and has uniformly bounded diameter in $\mathcal{C}(S)$. If $\|\mathbf{t}\|_{\boldsymbol{\alpha}} = 0$ then $\operatorname{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ contains the α_i and is contained in the 1-neighborhood of $\boldsymbol{\alpha}$ in $\mathcal{C}(S)$, and we are done.

Assume $\|\mathbf{t}\|_{\boldsymbol{\alpha}} > 0$, and let Y be a component of $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ with maximal area. Since $S(\mathbf{t}\cdot\boldsymbol{\alpha})$ has at most $\xi(S)$ components, we have $\operatorname{area}(Y) \geq \|\mathbf{t}\|_{\boldsymbol{\alpha}}^2/\xi(S)$. Note that Y cannot be an interval since $\|\mathbf{t}\|_{\boldsymbol{\alpha}} > 0$. To simplify the exposition, we first prove Lemma 5.1 when Y has genus at least 1, or at least 5 marked points — the case where Y is a sphere with 4 marked points shall be dealt with in Section 8B.

8A. ***Case 1: Y has genus at least 1 or at least 5 marked points.*** By Proposition 7.3 and Lemma 7.4, there exists an essential annulus A on Y with

$$\operatorname{width}(A) \geq \mathsf{W}_0\sqrt{\operatorname{area}(Y)} \geq \frac{\mathsf{W}_0\|\mathbf{t}\|_{\boldsymbol{\alpha}}}{\sqrt{\xi(S)}},$$

where $\mathsf{W}_0 = \mathsf{W}_0(\xi(Y))$. Let $\gamma \in \mathcal{C}(S)$ be the core curve of A. Setting

$$\mathsf{W} = \mathsf{W}(\xi(S)) := \min_{1 \leq k \leq \xi(S)} \frac{\mathsf{W}_0(k)}{\sqrt{\xi(S)}},$$

we have $\operatorname{width}(A) \geq \mathsf{W}\|\mathbf{t}\|_{\boldsymbol{\alpha}}$. Applying the Besicovitch Lemma [1952] (see Lemma $4.5\frac{1}{2}$ in [Gromov 1999] for a proof), we have

$$\operatorname{width}(A) \times \operatorname{length}(A) \leq \operatorname{area}(A),$$

where $\operatorname{length}(A)$ is the length of a shortest core curve on A. Since $\operatorname{area}(A)$ is at most $\operatorname{area}(S(\mathbf{t}\cdot\boldsymbol{\alpha})) = \|\mathbf{t}\|_{\boldsymbol{\alpha}}^2$, it follows that $l(\gamma) \leq \operatorname{length}(A) \leq \|\mathbf{t}\|_{\boldsymbol{\alpha}}/\mathsf{W}$.

Now let b be a geodesic representative of a curve $\beta \in \mathcal{C}(S)$ on $S(\mathbf{t}\cdot\boldsymbol{\alpha})$. Each essential intersection of b with A contributes at least $\operatorname{width}(A)$ to $\operatorname{length}(b)$, and so

$$\operatorname{width}(A) \times i(\gamma, \beta) \leq \operatorname{length}(b) = l(\beta).$$

Combining the above with Proposition 7.1, we deduce

$$i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) \leq \sqrt{2}\, l(\gamma) \leq \frac{\sqrt{2}\|\mathbf{t}\|_{\boldsymbol{\alpha}}}{\mathsf{W}} \quad \text{and} \quad i(\gamma, \beta) \leq \frac{l(\beta)}{\operatorname{width}(A)} \leq \frac{i(\mathbf{t}\cdot\boldsymbol{\alpha}, \beta)}{\mathsf{W}\|\mathbf{t}\|_{\boldsymbol{\alpha}}}.$$

Set $\mathsf{L}_0 = \sqrt{2}/\mathsf{W}$. The curve γ satisfies $i(\mathbf{t}\cdot\boldsymbol{\alpha}, \gamma) \leq \mathsf{L}_0 \|\mathbf{t}\|_{\boldsymbol{\alpha}}$, and so short$(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L}) \neq \varnothing$ for all $\mathsf{L} \geq \mathsf{L}_0$. Furthermore, if $\beta \in$ short$(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})$ then

$$i(\gamma, \beta) \leq \frac{i(\mathbf{t}\cdot\boldsymbol{\alpha}, \beta)}{\mathsf{W}\|\mathbf{t}\|_{\boldsymbol{\alpha}}} \leq \frac{\mathsf{L}\|\mathbf{t}\|_{\boldsymbol{\alpha}}}{\mathsf{W}\|\mathbf{t}\|_{\boldsymbol{\alpha}}} = \frac{\mathsf{L}}{\mathsf{W}}.$$

Applying Lemma 2.2 and the triangle inequality gives

$$\operatorname{diam}(\operatorname{short}(\mathbf{t}\cdot\boldsymbol{\alpha}, \mathsf{L})) \leq 2\Big[2\log_2\Big(\frac{\mathsf{L}}{\mathsf{W}}\Big) + 2\Big] = 4\log_2 \mathsf{L} + \mathsf{k}_0,$$

where k_0 is a constant depending only on $\xi(S)$. □

8B. ***Case 2: Y is a sphere with 4 marked points.*** For our purposes, it suffices to find a wide annulus on a suitable double branched cover of Y. Identify Y with the quotient of the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ under a hyperelliptic involution $h : \mathbb{T}^2 \to \mathbb{T}^2$ given by $h(x, y) = (-x, -y)$ modulo $\mathbb{Z}^2$, so that the marked points coincide with the branch points. Metrize $\mathbb{T}^2$ by pulling back the singular Euclidean metric on Y. This metric can also be obtained by taking the preimages of the α_i contained in Y to $\mathbb{T}^2$ and then applying the construction as described in Section 4A. It follows that $\mathbb{T}^2$ enjoys the isoperimetric inequality stated in Lemma 7.4, and so applying Proposition 7.3 gives the following:

Lemma 8.1. *There exists an essential annulus on $\mathbb{T}^2$ of width at least* $\mathsf{W}'\|\mathbf{t}\|_{\boldsymbol{\alpha}}/\sqrt{\xi(S)}$ *for some universal constant* $\mathsf{W}' > 0$.

Remark 8.2. By following the proof of Proposition 7.3 in [Bowditch 2006b] for the case of the torus, one can take $\mathsf{W}' = \frac{1}{3\sqrt{2}}$.

Observe that $h(\tilde{\gamma})$ is homotopic to $\tilde{\gamma}$ for any simple closed curve $\tilde{\gamma}$ on $\mathbb{T}^2$. Thus, any simple closed on $\mathbb{T}^2$ descends to a simple closed curve on Y (up to homotopy).

Lemma 8.3. *Let A be an essential annulus on $\mathbb{T}^2$ with core curve $\tilde{\gamma}$. Let $\gamma \in \mathcal{C}(S)$ be the image of $\tilde{\gamma}$ on Y under the quotient map $H : \mathbb{T}^2 \to Y$. Then*

$$i(\gamma, \beta) \leq \frac{2i(\mathbf{t}\cdot\boldsymbol{\alpha}, \beta)}{\operatorname{width}(A)}$$

for all $\beta \in \mathcal{C}(S)$.

Proof. Recall that $\beta \cap Y$ is either a simple closed curve or a union of paths connecting marked points of Y. The preimage $H^{-1}(\beta)$ is a finite union of (not necessarily disjoint) essential curves on $\mathbb{T}^2$. By perturbing γ to an embedded curve which misses the marked points of Y, we see that each point of $\gamma \cap \beta$ lifts to exactly two points on $\mathbb{T}^2$ under H, and so

$$i(\gamma, \beta) = \frac{i(H^{-1}(\gamma), H^{-1}(\beta))}{2} \leq i(\tilde{\gamma}, H^{-1}(\beta)).$$

By observing that each intersection of $H^{-1}(\beta)$ with A contributes at least width(A) to its length, and applying Proposition 7.1, we deduce

$$\operatorname{width}(A) \times i(\tilde{\gamma}, H^{-1}(\beta)) \le l(H^{-1}(\beta)) = 2l(\beta \cap Y) \le 2l(\beta) \le 2i(\mathbf{t} \cdot \boldsymbol{\alpha}, \beta).$$

The result follows. $\square$

We may use the previous lemmas and argue as in Case 1 to bound the diameter of short$(\mathbf{t} \cdot \boldsymbol{\alpha}, \mathsf{L})$. Finally, short$(\mathbf{t} \cdot \boldsymbol{\alpha}, \mathsf{L})$ is nonempty since $\boldsymbol{\alpha}$ does not fill S.

Acknowledgements

I am grateful to both Brian Bowditch and Saul Schleimer for many helpful discussions and suggestions. I also thank Saul Schleimer and Marc Lackenby for thorough comments on my PhD thesis [Tang 2013], in which the contents of this paper are completely contained. Many thanks to the anonymous referee for providing several suggestions to improve the exposition of this paper. This work was supported by a Warwick Postgraduate Research Scholarship.

References

[Alonso et al. 1991] J. M. Alonso, T. Brady, D. Cooper, V. Ferlini, M. Lustig, M. Mihalik, M. Shapiro, and H. Short, "Notes on word hyperbolic groups", pp. 3–63 in *Group theory from a geometrical viewpoint* (Trieste, 1990), edited by E. Ghys et al., World Scientific, River Edge, NJ, 1991. MR Zbl

[Aougab 2013] T. Aougab, "Uniform hyperbolicity of the graphs of curves", *Geom. Topol.* **17**:5 (2013), 2855–2875. MR Zbl

[Besicovitch 1952] A. S. Besicovitch, "On two problems of Loewner", *J. London Math. Soc.* **27** (1952), 141–144. MR Zbl

[Bowditch 2006a] B. H. Bowditch, *A course on geometric group theory*, MSJ Memoirs **16**, Mathematical Soc. Japan, Tokyo, 2006. MR Zbl

[Bowditch 2006b] B. H. Bowditch, "Intersection numbers and the hyperbolicity of the curve complex", *J. Reine Angew. Math.* **598** (2006), 105–129. MR Zbl

[Bowditch 2014] B. H. Bowditch, "Uniform hyperbolicity of the curve graphs", *Pacific J. Math.* **269**:2 (2014), 269–280. MR Zbl

[Bridson and Haefliger 1999] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren Math. Wissenschaften **319**, Springer, Berlin, 1999. MR Zbl

[Clay et al. 2014] M. Clay, K. Rafi, and S. Schleimer, "Uniform hyperbolicity of the curve graph via surgery sequences", *Algebr. Geom. Topol.* **14**:6 (2014), 3325–3344. MR Zbl

[Gromov 1987] M. Gromov, "Hyperbolic groups", pp. 75–263 in *Essays in group theory*, edited by S. M. Gersten, Math. Sci. Res. Inst. Publ. **8**, Springer, New York, 1987. MR Zbl

[Gromov 1999] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics **152**, Birkhäuser, Boston, 1999. MR Zbl

[Harvey 1981] W. J. Harvey, "Boundary structure of the modular group", pp. 245–251 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, 1981. MR Zbl

[Hempel 2001] J. Hempel, "3-manifolds as viewed from the curve complex", *Topology* **40**:3 (2001), 631–657. MR Zbl

[Hensel et al. 2015] S. Hensel, P. Przytycki, and R. C. H. Webb, "1-slim triangles and uniform hyperbolicity for arc graphs and curve graphs", *J. Eur. Math. Soc.* **17**:4 (2015), 755–762. MR Zbl

[Kobayashi 1988] T. Kobayashi, "Heights of simple loops and pseudo-Anosov homeomorphisms", pp. 327–338 in *Braids* (Santa Cruz, CA, 1986), edited by J. S. Birman and A. Libgober, Contemp. Math. **78**, Amer. Math. Soc., Providence, RI, 1988. MR Zbl

[Masur and Minsky 1999] H. A. Masur and Y. N. Minsky, "Geometry of the complex of curves, I: Hyperbolicity", *Invent. Math.* **138**:1 (1999), 103–149. MR Zbl

[Rafi and Schleimer 2009] K. Rafi and S. Schleimer, "Covers and the curve complex", *Geom. Topol.* **13**:4 (2009), 2141–2162. MR Zbl

[Schleimer 2005] S. Schleimer, "Notes on the complex of curves", lecture notes, Mathematics Institute, Warwick, 2005, Available at http://homepages.warwick.ac.uk/~masgar/Maths/notes.pdf.

[Tang 2012] R. Tang, "The curve complex and covers via hyperbolic 3-manifolds", *Geom. Dedicata* **161** (2012), 233–237. MR Zbl

[Tang 2013] R. Tang, *Covering maps and hulls in the curve complex*, Ph.D. thesis, University of Warwick, 2013, Available at http://search.proquest.com/docview/1685021303.

Received November 6, 2014. Revised June 11, 2016.

ROBERT TANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OKLAHOMA
NORMAN, OK 73019
UNITED STATES

rtang@math.ou.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 1, 2017
dx.doi.org/10.2140/pjm.2017.291.241

THE LOCAL GINZBURG–RALLIS MODEL OVER THE COMPLEX FIELD

CHEN WAN

We consider the local Ginzburg–Rallis model over the complex field. We show that the multiplicity is always 1 for a majority of generic representations. We also have partial results on the real case for general generic representations. This is a continuation of our previous work in which we considered the p-adic case and the real case for tempered representations.

1. Introduction and main result

This paper is a continuation of [Wan 2016a; 2016b]. For an overview of the Ginzburg–Rallis model, see Section 1 of [Wan 2016a]. We recall from there the definition of the Ginzburg–Rallis models and conjectures.

Let F be a local field (p-adic or archimedean), D be the unique quaternion algebra over F if $F \neq \mathbb{C}$. Take $P = P_{2,2,2} = MU$ to be the standard parabolic subgroup of $G = \mathrm{GL}_6$ whose Levi part M is isomorphic to $\mathrm{GL}_2 \times \mathrm{GL}_2 \times \mathrm{GL}_2$, and whose unipotent radical U consists of elements of the form

$$u = u(X, Y, Z) := \begin{pmatrix} I_2 & X & Z \\ 0 & I_2 & Y \\ 0 & 0 & I_2 \end{pmatrix}. \tag{1-1}$$

We define a character ξ on $U(F)$ by

$$\xi(u(X, Y, Z)) := \psi(\mathrm{tr}(X) + \mathrm{tr}(Y)), \tag{1-2}$$

where ψ is a nontrivial additive character on F. It's clear that the stabilizer of ξ is the diagonal embedding of $\mathrm{GL}_2(F)$ into $M(F)$, which is denoted by $H_0(F)$. For a given character χ of $F^\times$, one induces a one dimensional representation ω of $H_0(F)$ given by $\omega(h) := \chi(\det(h))$. We can extend the character ξ to the semidirect product

$$H(F) := H_0(F) \ltimes U(F) \tag{1-3}$$

by making it trivial on $H_0(F)$. Similarly we can extend the character ω to $H(F)$. It follows that the one dimensional representation $\omega \otimes \xi$ of $H(F)$ is well defined. The

MSC2010: 22E50.
Keywords: representations of linear algebraic groups over archimedean local field, multiplicity one on Vogan packet.

pair (G, H) is the Ginzburg–Rallis model, introduced in [Ginzburg and Rallis 2000]. Let π be an irreducible admissible representation of $G(F)$ with central character χ^2, we are interested in the Hom space $\mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)$, the dimension of which is denoted by $m(\pi)$ and is called the multiplicity.

On the other hand, if $F \neq \mathbb{C}$, define $G_D = \mathrm{GL}_3(D)$. Similarly we can define U_D, $H_{0,D}$ and H_D. We also define the character $\omega_D \otimes \xi_D$ on $H_D(F)$ in the same way except that the trace in the definition of ξ is replaced by the reduced trace of the quaternion algebra D and the determinant in the definition of ω is replaced by the reduced norm of the quaternion algebra D. Then for an irreducible admissible representation π_D of $G_D(F)$ with central character χ^2, we can also talk about the Hom space $\mathrm{Hom}_{H_D(F)}(\pi_D, \omega_D \otimes \xi_D)$, whose dimension is denoted by $m(\pi_D)$.

The purpose of this paper is to study the multiplicity $m(\pi)$ and $m(\pi_D)$. First, it was proved by C.-F. Nien [2006] over a p-adic local field, and by D. Jiang, B. Sun and C. Zhu in [Jiang et al. 2011] for an archimedean local field that both multiplicities are less than or equal to 1: $m(\pi), m(\pi_D) \leq 1$. In other word, the pairs (G, H) and (G_D, H_D) are Gelfand pairs. In this paper, we are interested in the relation between $m(\pi)$ and $m(\pi_D)$ under the local Jacquet–Langlands correspondence established in [Deligne et al. 1984]. The local conjecture has been expected since the work of [Ginzburg and Rallis 2000], and was first discussed in detail by Jiang [2008].

Conjecture 1.1 [Jiang 2008]. *For any irreducible admissible representation π of* $\mathrm{GL}_6(F)$, *let π_D be the local Jacquet–Langlands correspondence of π to* $\mathrm{GL}_3(D)$ *if it exists, and zero otherwise. In particular, π_D is always* 0 *if $F = \mathbb{C}$. We still assume that the central character of π is χ^2. Then the following identity*:

$$m(\pi) + m(\pi_D) = 1 \tag{1-4}$$

holds for all irreducible generic representations π of $\mathrm{GL}_6(F)$.

Note that the assertion in Conjecture 1.1 can be formulated in terms of Vogan packets and pure inner forms of PGL_6. We refer to [Wan 2016a] for discussion.

Another aspect of the local conjecture is the epsilon dichotomy conjecture, which relates the multiplicity to the central value of the exterior cube epsilon factors. It can be stated as follows:

Conjecture 1.2. *With the same assumptions as in Conjecture 1.1, assume that the central character of π is trivial. Then we have*

$$m(\pi) = 1 \Longleftrightarrow \epsilon\left(\tfrac{1}{2}, \pi, \textstyle\bigwedge^3\right) = 1, \quad m(\pi) = 0 \Longleftrightarrow \epsilon\left(\tfrac{1}{2}, \pi, \textstyle\bigwedge^3\right) = -1.$$

In this paper, we always fix a Haar measure dx on F and an additive character ψ such that the Haar measure is self-dual for Fourier transform with respect to ψ. We use such dx and ψ in the definition of the ϵ factor. For simplicity, we will write the epsilon factor as $\epsilon(s, \pi, \rho)$ instead of $\epsilon(s, \pi, \rho, dx, \psi)$.

Remark 1.3. Conjecture 1.2 can also be formulated for general representations with nontrivial central character. To be specific, as in Conjecture 1.1, assume the central character of π is χ^2. Then the epsilon dichotomy conjecture for π becomes

$$m(\pi)=1 \Longleftrightarrow \epsilon\big(\tfrac{1}{2},\pi,\textstyle\bigwedge^3\otimes\chi^{-1}\big)=1, \quad m(\pi)=0 \Longleftrightarrow \epsilon\big(\tfrac{1}{2},\pi,\textstyle\bigwedge^3\otimes\chi^{-1}\big)=-1.$$

Here, $\epsilon\big(s,\pi,\bigwedge^3\otimes\chi^{-1}\big)$ is the epsilon factor of $\big(\bigwedge^3\phi_\pi\big)\otimes\chi^{-1}$ (*not* $\bigwedge^3(\phi_\pi\otimes\chi^{-1})$), where ϕ_π is the Langlands parameter of π. The proof of the epsilon dichotomy conjecture for representations with nontrivial central characters is the same as the trivial central character case. Hence for simplicity, in this paper, we will only consider the trivial central character case for the epsilon dichotomy conjecture. All our results can be easily extended to the nontrivial central character case.

In the previous papers [Wan 2016a; 2016b], we prove Conjecture 1.1 for the case that F is a p-adic local field or $\mathbb{R}$ and π is an irreducible tempered representation of $\mathrm{GL}_6(F)$. In [Wan 2016b], we also prove Conjecture 1.2 for the case $F=\mathbb{R}$ and π is tempered, together with the case when F is p-adic and π is tempered but not a discrete series or a parabolic induction of a discrete series of $\mathrm{GL}_4(F)\times\mathrm{GL}_2(F)$.

In this paper, we consider the case when $F=\mathbb{C}$. In this case, by the Langlands classification, any generic representation π is a principal series. In other words, let $B=M_0U_0$ be the Borel subgroup consisting of all the lower triangular matrices; here $M_0=(\mathrm{GL}_1)^6$ is just the diagonal matrix. Then π is of the form $I_B^G(\chi)$, where $\chi=\bigotimes_{i=1}^6\chi_i$ is a character on $M_0(F)$ and I_B^G is the normalized parabolic induction. For $1\le i\le 6$, we can find a unitary character σ_i and some real number $s_i\in\mathbb{R}$ such that $\chi_i=\sigma_i|\cdot|^{s_i}$. Without loss of generality, we assume that $s_i\le s_j$ for any $i\ge j$. Then if we combine those representations with the same exponents s_i, we can find a parabolic subgroup $Q=LU_Q$ containing B with $L=\times_{i=1}^k\mathrm{GL}_{n_i}$, a representation $\tau=\bigotimes_{i=1}^k\tau_i|\cdot|^{t_i}$ of $L(F)$, where τ_i are all tempered and the exponents t_i are strictly increasing (i.e., $t_1<t_2<\cdots<t_k$) such that $\pi=I_Q^G(\tau)$. On the other hand, we can also write π as $I_{\bar P}^G(\pi_0)$ with $\pi_0=\pi_1\otimes\pi_2\otimes\pi_3$, where π_i is the parabolic induction of $\chi_{2i-1}\otimes\chi_{2i}$. Here we want the representation to be induced from $\bar P$ instead of P because later in Sections 5 and 6, we would like to integrate the elements of the induced representation over the unipotent subgroup $U(F)$.

Theorem 1.4. *Assume that $F=\mathbb{C}$, with the same assumptions as in Conjecture 1.1 and with the notation above. Then we have the following*:

(1) *If $\bar P\subset Q$, Conjectures 1.1 and 1.2 hold. In particular, both conjectures hold for the tempered representations.*

(2) *If $Q\subsetneq\bar P$ and if π_0 satisfies the condition* (40) *in [Loke 2001], Conjectures 1.1 and 1.2 hold.*

There are two main ingredients in our proof. First we deal with the tempered representations. The idea is to construct an explicit element inside the Hom space given by integrating the matrix coefficient. Then we show that the nonvanishing property of this element is invariant under parabolic induction, which allows us to reduce to the torus case which is trivial. This idea already appears in [Wan 2016b] for the case when $F=\mathbb{R}$.

Then for general generic representations, we use the open orbit method to reduce our problems to the tempered case or the trilinear GL_2 model case. To be specific, if $\bar{P}\subset Q$, by applying the open orbit method, we can reduce to the model related to the Levi subgroup L. Then after twisting τ by some characters, we only need to deal with the tempered case which has already been proved. If $Q\subsetneq \bar{P}$, by applying the open orbit method, we reduce ourselves to the trilinear GL_2 model case. Then by applying the work of Loke [2001], we can prove our result. The extra condition in part (2) of Theorem 1.4 also comes from the same work.

It is worth mentioning that in Theorem 1.4(2), the requirements we made for the parabolic subgroup Q force some types of generalized Jacquet integrals to be absolutely convergent; this allows us to apply the open orbit method. If one can prove such integrals have holomorphic continuation, one can actually remove this restraint. This will be discussed in Section 7.

Finally, the open orbit method we use here can also be applied to the case when $F=\mathbb{R}$; this will gives us partial results about Conjecture 1.1 and Conjecture 1.2 for general generic representations. To be specific, let π be an irreducible generic representation of $G(F)$ with central character χ^2. By the Langlands classification, there is a parabolic subgroup $Q=LU_Q$ containing the lower Borel subgroup and an essential tempered representation $\tau=\bigotimes_{i=1}^k \tau_i|\cdot|^{s_i}$ of $L(F)$ with τ_i tempered, $s_i\in\mathbb{R}$ and $s_1<s_2<\cdots<s_k$ such that $\pi=I_Q^G(\tau)$. We say Q is nice if $Q\subset\bar{P}$ or $\bar{P}\subset Q$.

Theorem 1.5. *Let the notation be as above.*

(1) *If* $\pi_D=0$ *and* Q *is nice, then Conjectures 1.1 and 1.2 hold.*

(2) *If* $\pi_D\neq 0$, *we have*

$$m(\pi)+m(\pi_D)\geq 1,$$

and if moreover if the central character of π *is trivial* (*as in* Conjecture 1.2), *we have*

$$\epsilon\big(\tfrac{1}{2},\pi,\textstyle\bigwedge^3\big)=1\Rightarrow m(\pi)=1;\quad m(\pi)=0\Rightarrow \epsilon\big(\tfrac{1}{2},\pi,\textstyle\bigwedge^3\big)=-1.$$

As in the complex case, the assumption on Q can be removed if we can prove the holomorphic continuation of certain generalized Jacquet integrals. This will also be discussed in Section 7.

The paper is organized as follows: In Section 2, we review a well know result of the intertwining operator which is due to Harish-Chandra. We will also give a brief overview of the open orbit method which will be used in later sections. In Section 3, we show that for $F = \mathbb{C}$, Conjecture 1.1 implies Conjecture 1.2. In Section 4, we prove Theorem 1.4 for tempered representations. Then in Section 5, we prove it for general cases. In Section 6, we discuss the case for $F = \mathbb{R}$. In Section 7, we talk about how to remove the assumptions on Q based on the results on the holomorphic continuation of the generalized Jacquet integral due to Raul Gomez [$\geq$ 2017].

2. Preliminaries

2A. ***The intertwining operator.*** For every connected reductive algebraic group G defined over F, let A_G be the maximal split center of G and Z_G be the center of G. We denote by $X(G)$ the group of F-rational characters of G. Define $\mathfrak{a}_G = \mathrm{Hom}(X(G), \mathbb{R})$, and let $\mathfrak{a}_G^* = X(G) \otimes_{\mathbb{Z}} \mathbb{R}$ be the dual of $\mathfrak{a}_G$. We define a homomorphism $H_G : G(F) \to \mathfrak{a}_G$ by $H_G(g)(\chi) = \log(|\chi(g)|_F)$ for every $g \in G(F)$ and $\chi \in X(G)$.

Given a parabolic subgroup $P = MU$ of G and an admissible representation (τ, V_τ) of $M(F)$, let K be a maximal compact subgroup of $G(F)$ in good position with respect to M. Let $(I_P^G(\tau), I_P^G(V_\tau))$ be the normalized parabolic induced representation: $I_P^G(V_\tau)$ consists of smooth functions $e : G(F) \to V_\tau$ such that

$$e(mug) = \delta_P(m)^{1/2}\tau(m)e(g), \quad m \in M(F), \quad u \in U(F), \quad g \in G(F),$$

and the $G(F)$-action is just the right translation.

For $\lambda \in \mathfrak{a}_M^* \otimes_{\mathbb{R}} \mathbb{C}$, let τ_λ be the unramified twist of τ, i.e.,

$$\tau_\lambda(m) = \exp(\lambda(H_M(m)))\tau(m),$$

and let $I_P^G(\tau_\lambda)$ be the induced representation. By the Iwasawa decomposition, every function $e \in I_P^G(\tau_\lambda)$ is determined by its restriction on K, and that space is invariant under the unramified twist, i.e., for any λ, we can realize the representation $I_P^G(\tau_\lambda)$ on the space $I_{K\cap P}^K(\tau_K)$ which consists of functions $e_K : K \to V_\tau$ such that

$$e(mug) = \delta_P(m)^{1/2}\tau(m)e(g), \quad m \in M(F) \cap K, \quad u \in U(F) \cap K, \quad g \in K.$$

Here τ_K is the restriction of the representation τ to the group K.

Now we define the intertwining operator. For a Levi subgroup M of G, $P = MU$, $P' = MU' \in \mathcal{P}(M)$, and $\lambda \in \mathfrak{a}_M^* \otimes_{\mathbb{R}} \mathbb{C}$, define the intertwining operator

$$J_{P'|P}(\tau_\lambda) : I_P^G(V_\tau) \to I_{P'}^G(V_\tau), \quad J_{P'|P}(\tau_\lambda)(e)(g) = \int_{(U(F)\cap U'(F))\backslash U'(F)} e(ug)\,du.$$

In general, the integral above is not absolutely convergent. But it is absolutely convergent when $\mathrm{Re}(\lambda)$ lies inside a positive cone, and it is $G(F)$-equivariant. By restricting to K, we can view $J_{P'|P}(\tau_\lambda)$ as a homomorphism from $I_{K\cap P}^K(V_{\tau_K})$ to $I_{K\cap P'}^K(V_{\tau_K})$. In general, $J_{P'|P}(\tau_\lambda)$ can be meromorphically continued to a function on $\mathfrak{a}_M^* \otimes_{\mathbb{R}} \mathbb{C}$. Moreover, if we assume that τ is tempered, we have the following proposition which is due to Harish-Chandra. The proof of the proposition can be found in Proposition IV.2.1 of [Waldspurger 2003].

Proposition 2.1. *With the notation above, if τ is tempered, then the intertwining operator $J_{P'|P}$ is absolutely convergent for all $\lambda \in \mathfrak{a}_M^* \otimes_{\mathbb{R}} \mathbb{C}$ with $\langle \mathrm{Re}(\lambda), \check{\alpha} \rangle > 0$ for every $\alpha \in \Sigma(P) \cap \Sigma(\bar{P}')$. Here $\Sigma(P)$ is the subset of the roots of A_M that are positive with respect to P.*

We will use this proposition in later sections to show some generalized Jacquet integrals are absolutely convergent.

Finally, assume π is a unitary representation of $G(F)$. Let $\mathrm{End}(\pi)$ be the space of continuous endomorphisms of π. We define the norm on $\mathrm{End}(\pi)$ to be $\|T\| = \sup_{e \in \pi,\, |e|=1} |Te|$. Then it becomes a Banach space. It is also a continuous representation of $G(F) \times G(F)$ under the left and right translations. Let $\mathrm{End}(\pi)^\infty$ be the subspace of smooth vectors. We can define a locally convex topology on $\mathrm{End}(\pi)^\infty$ via the seminorms

$$\|T\|_{u,v} = \|\pi(u)T\pi(v)\|, \quad u, v \in \mathcal{U}(\mathfrak{g}), \quad T \in \mathrm{End}(\pi)^\infty.$$

Here $\mathcal{U}(\mathfrak{g})$ is the universal enveloping algebra. This makes it a Fréchet space.

2B. *The open orbit method.* In this section we will give a brief overview of the open orbit method. The purpose of this method is to study the distinction of induced representations; it is an application of the geometric lemma due to Bernstein and Zelevinsky [1977]. Let G be a connected reductive group defined over F, and $H \subset G$ be a closed subgroup such that $X = H \backslash G$ is a spherical variety of G (i.e., the Borel subgroup has an open orbit). Let $P = MU$ be a parabolic subgroup of G and (τ, V_τ) be an irreducible admissible representation of $M(F)$. We want to study the Hom space $\mathrm{Hom}_{H(F)}(I_P^G(\tau), \chi)$, where χ is some character of $H(F)$. We say $(\pi, V_\pi) = (I_P^G(\tau), I_P^G(V_\tau))$ is (H, χ)-distinguished (or just H-distinguished if χ is trivial) if the Hom space is nonzero. For simplicity, we assume that χ is trivial.

The geometric lemma [Bernstein and Zelevinsky 1977]. *There is an ordering $\{P(F)y_i H(F)\}_{i=1}^N$ on the double coset $H(F) \backslash G(F)/P(F)$ such that*

$$Y_i = \bigcup_{j=1}^{i} P(F)y_i H(F)$$

is open in $G(F)$ for any $1 \leq i \leq N$.

With the filtration above, for $1 \leq i \leq N$, define

$$V_i = \{f \in I_P^G(V_\tau) \mid \mathrm{supp}(f) \subset Y_i\}.$$

Then we have $V_1 \subset V_2 \subset \cdots \subset V_N = V_\pi$ and V_i is $H(F)$-invariant for all i. In particular, this implies that if $I_P^G(\tau)$ is H-distinguished, there exists i such that $\mathrm{Hom}_{H(F)}(V_i/V_{i-1}, \chi) \neq 0$ (here $V_0 = \{0\}$). Moreover, for any $1 \leq i \leq N$, it is easy to see that the map

$$f \in V_i \mapsto \phi_f(h) := f(y_i h)$$

is an isomorphism between V_i/V_{i-1} and $\mathrm{ind}_{H_i}^H(\delta_P^{1/2}\tau^{y_i}|_{H_i})$ ($\mathrm{ind}_{H_i}^H$ is the compact induction). Here $H_i = H(F) \cap y_i^{-1}P(F)y_i = y_i^{-1}P_i y_i$, with $P_i = P(F) \cap y_i H(F) y_i^{-1}$. By applying the reciprocity law, we have a necessary condition for $I_P^G(\tau)$ to be H-distinguished.

Proposition 2.2. *If $I_P^G(\tau)$ is H-distinguished, then there exists i such that τ is $(P_i, \delta_{P_i}\delta_P^{1/2})$-distinguished. Here we view τ as a representation of $P(F)$ by making it trivial on $U(F)$.*

What we are interested is the opposite direction of the proposition above. In other words, we want to have some sufficient conditions for $I_P^G(\tau)$ to be H-distinguished in terms of V_i/V_{i-1}. These are known as the open orbit method and the closed orbit method. For our purposes, we only consider the open orbit method.

Assume that τ is $(P_1, \delta_{P_1}\delta_P^{1/2})$-distinguished, we want to show that π is H-distinguished. For simplicity, assume that $H(F)P(F)$ is open in $G(F)$ and $y_1 = 1$. Choose a nonzero element l_0 in the Hom space for τ; it gives a nonzero element l in $\mathrm{Hom}_{H(F)}(V_1, 1)$ by integrating l_0 over $H_1(F)\backslash H(F)$. Then we would like to extend this integral to V_π, which will gives us a nonzero element in $\mathrm{Hom}_H(F)(V_\pi, 1)$. However, the integral will not be absolutely convergent in general; one needs to show that it has holomorphic continuation. In our case, the integral over $H(F)/H_1(F)$ will be some generalized Jacquet integral. In Sections 5 and 6, we will use Proposition 2.1 to show that the integral is absolutely convergent for some π with positive exponents. This will prove Theorem 1.4 and Theorem 1.5. Then in Section 7, we will talk about how to remove the restraints on the exponents by applying R. Gomez's result on the holomorphic continuation of generalized Jacquet integrals.

3. The relation between Conjectures 1.1 and 1.2

The goal of this section is to prove the following proposition:

Proposition 3.1. *If $F = \mathbb{C}$, then Conjecture 1.1 implies Conjecture 1.2.*

Proof. Since $F = \mathbb{C}$, π_D is always 0. Hence Conjecture 1.1 tells us that the multiplicity $m(\pi)$ is always 1. Therefore in order to prove Conjecture 1.2, it is enough to show that the epsilon factor $\epsilon\left(\frac{1}{2}, \pi, \bigwedge^3\right)$ equals 1 for any irreducible generic representations π of $\mathrm{GL}_6(F)$ with trivial central character.

By the Langlands classification, we can find a generic representation $\sigma = \sigma_1 \otimes \sigma_2$ of $\mathrm{GL}_5(F) \times \mathrm{GL}_1(F)$ such that π is the parabolic induction of σ. Let ϕ be the

Langlands parameter of π and ϕ_i be the Langlands parameter of σ_i for $i=1,2$. We have $\phi=\phi_1\oplus\phi_2$. This implies

$$\bigwedge^3(\phi)=\bigwedge^3(\phi_1\oplus\phi_2)=\bigwedge^3(\phi_1)\oplus(\bigwedge^2(\phi_1)\otimes\phi_2).$$

Since the central character of π is trivial, $\det(\phi)=\det(\phi_1)\otimes\det(\phi_2)=1$. Therefore $(\bigwedge^3(\phi_1))^\vee=\bigwedge^2(\phi_1)\otimes\det(\phi_1)^{-1}=\bigwedge^2(\phi_1)\otimes\det(\phi_2)=\bigwedge^2(\phi_1)\otimes\phi_2$, hence

$$\epsilon\bigl(\tfrac{1}{2},\pi,\bigwedge^3\bigr)=\det(\bigwedge^3(\phi_1))(-1)=(\det(\phi_1))^6(-1)=1.$$

This finishes the proof of the proposition. □

4. The tempered case

In this section, we prove our main theorem for the tempered case; the method is very similar to the case $F=\mathbb{R}$ we proved in [Wan 2016b]. Let π be a tempered representation of $G=\mathrm{GL}_6(F)$ with central character χ^2. Our goal is to show that $m(\pi)=1$. Since we already know that $m(\pi)\le 1$, it is enough to show that

$$m(\pi)\neq 0. \tag{4-1}$$

For all $T\in\mathrm{End}(\pi)^\infty$, define

$$\mathcal{L}_\pi(T)=\int^*_{Z_H(F)\backslash H(F)}\mathrm{tr}(\pi(h^{-1})T)\omega\otimes\xi(h)\,dh.$$

Here $\int^*_{Z_H(F)\backslash H(F)}$ is the normalized integral defined in Proposition 5.1 of [Wan 2016b]. Note the arguments in the loc. cit. is for the case when $F=\mathbb{R}$, but they also work for $F=\mathbb{C}$. For details, see the proof of Proposition 6.1.1 of [Wan 2017]. By Lemma 5.2 of [Wan 2016b] or Lemma 6.1.2 of [Wan 2017], for any $h,h'\in H(F)$,

$$\mathcal{L}_\pi(\pi(h)T\pi(h'))=\omega\otimes\xi(hh')\mathcal{L}_\pi(T). \tag{4-2}$$

For $e,e'\in\pi$, define $T_{e,e'}\in\mathrm{End}(\pi)^\infty$ by $e_0\in\pi\mapsto(e_0,e')e$. Set $\mathcal{L}_\pi(e,e')=\mathcal{L}_\pi(T_{e,e'})$. Then

$$\mathcal{L}_\pi(e,e')=\int^*_{Z_H(F)\backslash H(F)}(e,\pi(h)e')\omega\otimes\xi(h)\,dh.$$

If we fix e', by (4-2), the map $e\in\pi\to\mathcal{L}_\pi(e,e')$ belongs to $\mathrm{Hom}_H(\pi,\omega\otimes\xi)$. Since $\mathrm{Span}\{T_{e,e'}\mid e,e'\in\pi\}$ is dense in $\mathrm{End}(\pi)^\infty$, we have $\mathcal{L}_\pi\neq 0\Rightarrow m(\pi)\neq 0$. Hence in order to show the multiplicity $m(\pi)$ is nonzero, it is enough to show that the operator $\mathcal{L}_\pi$ is nonzero.

Since we are in the complex case, only $\mathrm{GL}_1(F)$ has discrete series; hence π is a principal series. Let $R=M_RU_R$ be a good minimal parabolic subgroup of G in the sense that RH is Zariski open in G. The existence of such R is proved in Proposition 4.2 of [Wan 2016b]. It is also proved in the same proposition that for all such R, we have $H_R := H \cap R = Z_G$. Hence the reduced model associated to R is just (M_R, Z_G). Since π is a principal series, there is a unitary character τ of $M_R(F)$ such that $\pi = I_R^G(\tau)$. For $T_0 \in \mathrm{End}(\tau)^\infty$, define

$$\mathcal{L}_\tau(T_0) = \mathrm{tr}(T_0).$$

By Proposition 5.9 of [Wan 2016b], the nonvanishing property of $\mathcal{L}_\pi$ is invariant under the parabolic induction, hence we have $\mathcal{L}_\pi \neq 0 \iff \mathcal{L}_\tau \neq 0$. Here the arguments in the loc. cit. is for the case when $F = \mathbb{R}$, but they also work for $F = \mathbb{C}$. Since $\mathcal{L}_\tau$ is obviously nonzero, we have $\mathcal{L}_\pi \neq 0$. This proves $m(\pi) \neq 0$ and hence finishes the proof of Theorem 1.4 for tempered representations.

5. The proof of Theorem 1.4

5A. ***The case when*** $\bar{P} \subset Q$**.** In this section, we prove the first part of Theorem 1.4. In other words, we assume that $\bar{P} \subset Q$. Then there are four possibilities for Q: type (6), type (4, 2), type (2, 4) or type (2, 2, 2). The idea is to first reduce our problem to the reduced model $(L, H \cap Q)$ by the open orbit method, then reduce it to the tempered case which was considered in the previous section.

If $Q = G$ **is of type** (6), by twisting π by some characters, we can assume that π is tempered. Note that twisting by characters will not change the multiplicities. Then by applying the result in the last section, we know that $m(\pi) \neq 0$ and this proves Theorem 1.4.

If Q **is of type** (4, 2), then $L(F) = \mathrm{GL}_4(F) \times \mathrm{GL}_2(F)$ and $H_Q(F) = H(F) \cap Q(F)$ is of the form

$$H_Q(F) = H_0(F) \ltimes U_{0,Q}(F),$$

where

$$U_{0,Q}(F) = \left\{ u = u(X) := \begin{pmatrix} 1 & X & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; X \in M_2(F) \right\}.$$

The restriction of the character ξ on $U_{0,Q}(F)$ is just $\xi(u(X)) = \psi(\mathrm{tr}(X))$ and the character ω on $H_0(F)$ is defined as usual. The model (L, H_Q) is the middle model introduced in [Wan 2016a]; it can be understood as the model between the Ginzburg–Rallis model and the trilinear GL_2 model. By the definition of Q, π is of the form $I_Q^G(\tau_1|\cdot|^{t_1} \otimes \tau_2|\cdot|^{t_2})$, where τ_1, τ_2 are tempered and $t_1 < t_2$. Hence any element $f \in \pi$ is a smooth function $f : G(F) \to \tau = \tau_1|\cdot|^{t_1} \otimes \tau_2|\cdot|^{t_2}$ such that

$$f(lug) = \delta_Q(l)^{1/2}\tau(l) f(g) \tag{5-1}$$

for all $l \in L(F)$, $u \in U_Q(F)$ and $g \in G(F)$. Here we use the letters π, σ, τ to denote both the representations and the underlying vector spaces. Let $\bar{Q} = LU_{\bar{Q}}$ be the opposite parabolic subgroup of Q. It is easy to see that $U_{\bar{Q}} \subset U$ and $U = U_{\bar{Q}} U_{0,Q}$. For any $f \in \pi$, define

$$J_Q(f) = \int_{U_{\bar{Q}}(F)} f(u)\xi^{-1}(u)\,du. \tag{5-2}$$

By Proposition 2.1 together with the assumption that $t_1 < t_2$, the integral above is absolutely convergent.

Proposition 5.1. (1) *For all $f \in \pi$, $u \in U_{\bar{Q}}(F)$ and $l \in H_Q(F)$, we have*

$$J_Q(\pi(u)f) = \xi(u)J(f) \tag{5-3}$$

and

$$J_Q(\pi(l)f) = \tau(l)J(f). \tag{5-4}$$

(2) *The function*

$$J_Q : \pi \to \tau, \quad f \to J_Q(f)$$

is surjective.

Proof. Part (1) follows from (5-1) and changing variables in the integral (5-2). For part (2), fix a function $\varphi \in C_c^\infty(U_{\bar{Q}}(F))$ such that $\int_{U_{\bar{Q}}(F)} \varphi(u)\psi^{-1}(u)\,du = 1$. For any $v \in \tau$, since $Q(F)U_{\bar{Q}}(F)$ is open in $G(F)$, the function

$$f(g) = \begin{cases} \delta_Q(l)^{1/2}\tau(l)\varphi(u)v & \text{if } g = u'lu \text{ with } l \in L(F), u \in U_{\bar{Q}}(F), u' \in U_Q(F), \\ 0 & \text{else} \end{cases}$$

lies inside π. Then we have

$$J_Q(f) = \int_{U_{\bar{Q}}(F)} f(u)\psi^{-1}(u)\,du = \int_{U_{\bar{Q}}(F)} \varphi(u)\psi^{-1}(u)v\,du = v.$$

This proves (2). □

We consider the Hom space $\mathrm{Hom}_{H_Q(F)}(\tau, (\omega \otimes \xi)|_{H_Q(F)})$ and let $m(\tau)$ be the dimension of that space. The following proposition tells us the relation between $m(\pi)$ and $m(\tau)$:

Proposition 5.2. $m(\tau) \neq 0 \Rightarrow m(\pi) \neq 0.$

Proof. If $m(\tau) \neq 0$, choose $0 \neq l_0 \in \mathrm{Hom}_{H_Q(F)}(\tau, (\omega\otimes\xi)|_{H_Q(F)})$. Define an operator l on π to be

$$l(f) = l_0(J_Q(f)).$$

Since $l_0 \neq 0$ and J_Q is surjective, we have $l \neq 0$. Hence we only need to show that $l \in \mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)$.

For $h \in H(F)$, we can write $h = h_1 u_1$ with $h_1 \in H_Q(F)$ and $u_1 \in U_{\bar{Q}}(F)$. By (5-3) and (5-4), we have

$$\begin{aligned} l(\pi(h)f) &= l_0(J_Q(\pi(h_1u_1)f)) &&= l_0(\tau(h_1)J_Q(\pi(u_1)f)) \\ &= \omega\otimes\xi(h_1)l_0(J_Q(\pi(u_1)f)) &&= \omega\otimes\xi(h_1)l_0(\xi(u_1)J_Q(f)) \\ &= \omega\otimes\xi(h)l_0(J_Q(f)) &&= \omega\otimes\xi(h)l(f). \end{aligned}$$

This implies $l \in \mathrm{Hom}_{H(F)}(\pi, \omega\otimes\xi)$ and finishes the proof of the proposition. $\square$

By the proposition above, we only need to show that $m(\tau) \neq 0$. It is easy to see that the multiplicity $m(\tau)$ is invariant under the unramified twist, hence we may assume that τ is tempered (note that originally τ is of the form $\tau_1|\cdot|^{t_1} \otimes \tau_2|\cdot|^{t_2}$ with τ_1 and τ_2 being tempered). Then by applying the argument in the previous section to the middle model case, we can show that the multiplicity $m(\tau)$ is always nonzero for all tempered representations τ. This proves Theorem 1.4.

If Q is of type $(2,4)$, the argument is the same as the $(4,2)$ case; we skip it here.

If Q is of type $(2,2,2)$, the argument is still similar to the $(4,2)$ case: we first reduce to the trilinear GL_2 model case by the open orbit method. Then after twisting by some characters we only need to consider the tempered case. Finally, by applying the argument in the previous section to the trilinear GL_2 model case, we can show that the multiplicity is nonzero and this proves Theorem 1.4. We skip the details here.

Now the proof of Theorem 1.4(1) is complete.

5B. ***The case when $Q \subsetneq \bar{P}$.*** In this section, we prove part (2) of Theorem 1.4. Recall that in Section 1 we assume that $\pi = I_B^G\big(\bigotimes_{i=1}^6 \chi_i\big)$, where B is the lower Borel subgroup, $\chi_i = \sigma_i|\cdot|^{s_i}$, σ_i are unitary characters, and s_i are real numbers with $s_1 \le s_2 \le \cdots \le s_6$. By the assumption $Q \subsetneq \bar{P}$, we have $s_2 < s_3$ and $s_4 < s_5$. Also as in Section 1, we write $\pi = I_{\bar{P}}^G(\pi_0)$, with $\pi_0 = \pi_1\otimes\pi_2\otimes\pi_3$ and π_i be the parabolic induction of $\chi_{2i-1}\otimes\chi_{2i}$. Then π consists of smooth functions $f \to \pi_0$ such that

$$f(mug) = \delta_{\bar{P}}(m)^{1/2}\pi_0(m)f(g) \tag{5-5}$$

for all $m \in M(F)$, $u \in \bar{U}(F)$ and $g \in G(F)$. We still want to apply the open orbit method. For $f \in \pi$, define

$$J(f) = \int_{U(F)} f(ug)\xi^{-1}(u)\,du. \tag{5-6}$$

By Proposition 2.1 together with the assumption on the exponents s_i, the integral above is absolutely convergent. Similarly as in the previous section, we can show

$$m(\pi_0) \neq 0 \Rightarrow m(\pi) \neq 0. \tag{5-7}$$

Here $m(\pi_0)$ is the multiplicity for the trilinear GL_2 model. In fact, for $0 \neq l_0 \in \mathrm{Hom}_{H_0(F)}(\pi_0, \omega)$. By a similar argument as in Proposition 5.2, we know that

$$l(f) := l_0(J(f))$$

is a nonzero element in $\mathrm{Hom}_{H(F)}(\pi, \omega \otimes \xi)$. This proves (5-7). Now by our assumption on π_0 together with the work by Loke [2001] for the trilinear GL_2 model, we know that $m(\pi_0) \neq 0$. This implies $m(\pi) \neq 0$ and finishes the proof of Theorem 1.4.

Remark 5.3. The assumption $Q \subsetneq \bar{P}$ is only used to make the generalized Jacquet integral $J(f)$ be absolutely convergent. Hence in general, if one can prove the holomorphic continuation of the generalized Jacquet integral $J(f)$, then the assumption $Q \subsetneq \bar{P}$ in Theorem 1.4(2) can be removed. This will be discussed in Section 7.

6. The proof of Theorem 1.5

In this section, by applying the open orbit method to the case when $F = \mathbb{R}$, we prove Theorem 1.5. Let π be an irreducible generic representation of $G(F)$ with central character χ^2. With the notation as in Section 1, there is a parabolic subgroup $Q = LU_Q$ containing the lower Borel subgroup and an essential tempered representation $\tau = \bigotimes_{i=1}^{k} \tau_i |\cdot|^{s_i}$ of $L(F)$ with τ_i tempered, $s_i \in \mathbb{R}$ and $s_1 < s_2 < \cdots < s_k$ such that $\pi = I_Q^G(\tau)$.

6A. ***The case when $\pi_D = 0$.*** In this section we assume that $\pi_D = 0$. Then by our assumptions in Theorem 1.5, Q is nice. If $Q \subset \bar{P}$, let $\pi_0 = I_{Q\cap M}^M(\tau)$. It is a generic representation of $M(F)$ and we have $\pi = I_{\bar{P}}^G(\pi_0)$. By the same argument as in Section 5B, we can show that

$$m(\pi_0) \neq 0 \Rightarrow m(\pi) \neq 0, \tag{6-1}$$

where $m(\pi_0)$ is the multiplicity of the trilinear GL_2 model. Since $\pi_D = 0$, the Jacquet–Langlands correspondence of π_0 from $M(F) = (\mathrm{GL}_2(F))^3$ to $(\mathrm{GL}_1(D))^3$ is zero. By applying the result for the trilinear GL_2 model in [Prasad 1990] and [Loke 2001], we have $m(\pi_0) = 1$. Combining with (6-1), we know $m(\pi) \neq 0$. Hence $m(\pi) = 1$, since we already know $m(\pi) \leq 1$. Therefore

$$m(\pi) + m(\pi_D) = m(\pi) = 1.$$

This proves Conjecture 1.1. For Conjecture 1.2, we only need to show that when $\pi_D = 0$, the epsilon factor $\epsilon\left(\frac{1}{2}, \pi, \bigwedge^3\right)$ is always 1. Since $\pi_D = 0$, by the local Jacquet–Langlands correspondence in [Deligne et al. 1984], π_0 is not an essential discrete series (i.e., discrete series twisted by characters), hence at least one of the π_i $(i = 1, 2, 3)$ is a principal series. Therefore we can find a generic representation

$\sigma = \sigma_1 \otimes \sigma_2$ of $\mathrm{GL}_5(F) \times \mathrm{GL}_1(F)$ such that π is the parabolic induction of σ. Then by the same argument as in Section 3, we can show that

$$\epsilon\left(\tfrac{1}{2}, \pi, \textstyle\bigwedge^3\right) = 1.$$

This finishes the proof of Conjecture 1.2.

If $\bar{P} \subset Q$, there are only four possibilities for Q: type (6), $(4,2)$, $(2,4)$ and $(2,2,2)$. If Q is type (6), by twisting π by some characters we can assume that π is tempered. Then both Conjectures 1.1 and 1.2 are proved in [Wan 2016b]. If Q is type $(4,2)$ or $(2,4)$, by the same argument as in Section 5A, we can reduce to the middle model case by the open orbit method. Then by twisting some characters, we only need to consider the tempered case which has already been proved in [Wan 2016b]. If Q is type $(2,2,2)$, the argument is similar except replacing the middle model by the trilinear GL_2 model.

Now the proof of Theorem 1.5(1) is complete.

6B. *The case when $\pi_D \neq 0$.* In this section we assume that $\pi_D \neq 0$. As a result, $\pi = I_{\bar{P}}^G(\pi_0)$ is the parabolic induction of some essential discrete series

$$\pi_0 = \pi_1|\cdot|^{s_1} \otimes \pi_2|\cdot|^{s_2} \otimes \pi_3|\cdot|^{s_3}$$

of $M(F)$, where the π_i are discrete series of $\mathrm{GL}_2(F)$ and s_i are real numbers. As usual, we assume that $s_1 \le s_2 \le s_3$. We can write π_D in the form $I_{\bar{P}_D}^{G_D}(\pi_{0,D})$, where $\pi_{0,D} = \pi_{1,D}|\cdot|^{s_1} \otimes \pi_{2,D}|\cdot|^{s_2} \otimes \pi_{3,D}|\cdot|^{s_3}$ is the Jacquet–Langlands correspondence of π_0 from $M(F)$ to $M_D(F)$. Let $m(\pi_0)$ (resp. $m(\pi_{0,D})$) be the multiplicity of the trilinear $\mathrm{GL}_2(F)$ (resp. $\mathrm{GL}_1(D)$) model.

Proposition 6.1. *With the notation above, in order to prove Theorem 1.5(2), it is enough to show that*

$$m(\pi_0) \neq 0 \Rightarrow m(\pi) \neq 0; \quad m(\pi_{0,D}) \neq 0 \Rightarrow m(\pi_D) \neq 0. \tag{6-2}$$

Proof. By Prasad's result for the trilinear GL_2 model, we have

$$m(\pi_0) + m(\pi_{0,D}) = 1. \tag{6-3}$$

Moreover, if we assume the central character of π_0 is trivial on $H_0(F)$, we have

$$m(\pi_0) = 1 \iff \epsilon\left(\tfrac{1}{2}, \pi_0\right) = 1; \quad m(\pi) = 0 \iff \epsilon\left(\tfrac{1}{2}, \pi_0\right) = -1. \tag{6-4}$$

Combining (6-2) and (6-3), we have $m(\pi) + m(\pi_D) \ge 1$; this proves the first part of Theorem 1.5(2). For the second part, assume that the central character of π is trivial. In Section 6.2 of [Wan 2016b], we proved that

$$\epsilon\left(\tfrac{1}{2}, \pi, \textstyle\bigwedge^3\right) = \epsilon\left(\tfrac{1}{2}, \pi_0\right). \tag{6-5}$$

Now if $\epsilon\big(\frac{1}{2},\pi,\bigwedge^3\big)=1$, by (6-5), we have $\epsilon\big(\frac{1}{2},\pi_0\big)=1$. Combining with (6-4), we have $m(\pi_0)=1$, therefore $m(\pi)=1$ by (6-2). On the other hand, if $m(\pi)=0$, by (6-2), we have $m(\pi_0)=0$. Combining with (6-4), we have $\epsilon\big(\frac{1}{2},\pi_0\big)=-1$, therefore $\epsilon\big(\frac{1}{2},\pi,\bigwedge^3\big)=-1$ by (6-5). This finishes the proof of Theorem 1.5(2). $\square$

By the proposition above, it is enough to prove (6-2). If $s_1=s_2=s_3$, by twisting π by some characters, we may assume that π is tempered (note that the multiplicities for both the Ginzburg–Rallis model and the trilinear GL_2 model are invariant under twisting by characters). Then the relation (6-2) has already been proved in Corollary 5.13 of [Wan 2016b]. In fact, in this case, we even have $m(\pi)=m(\pi_0)$ and $m(\pi_D)=m(\pi_{0,D})$.

If $s_1<s_2=s_3$, let $\pi_{2,3}$ be the parabolic induction of $\pi_2\otimes\pi_3$; it is a tempered representation of $\mathrm{GL}_4(F)$. We also know that π will be the parabolic induction of $\pi'=\pi_1|\cdot|^{s_1}\otimes\pi_{2,3}|\cdot|^{s_2}$. Let $m(\pi')$ be the multiplicity for the middle model. By applying the open orbit method as in Section 5A, we have

$$m(\pi')\neq 0\Rightarrow m(\pi)\neq 0.$$

Hence in order to prove $m(\pi_0)\neq 0\Rightarrow m(\pi)\neq 0$, it is enough to show that $m(\pi_0)\neq 0\Rightarrow m(\pi')\neq 0$. Again by twisting π' by some characters, we may assume that π' is tempered. Then by Corollary 5.13 of [Wan 2016b], we have $m(\pi_0)=m(\pi')$, which implies $m(\pi_0)\neq 0\Rightarrow m(\pi)\neq 0$. The proof of the quaternion version is similar. This proves (6-2).

If $s_1=s_2<s_3$, the argument is the same as the case above; we skip it here.

If $s_1<s_2<s_3$, (6-2) follows directly from the open orbit method as in Section 5A.

Now the proof of Theorem 1.5(2) is complete.

7. Holomorphic continuation of the generalized Jacquet integral

In the previous sections, we have already seen that the extra conditions on Q in Theorem 1.4(2) and Theorem 1.5(1) can be removed if the generalized Jacquet integral $J(f)$ defined in (5-6) has holomorphic continuation. In this section, we are going to remove the condition on Q based on the following hypothesis:

Hypothesis: The generalized Jacquet integrals have holomorphic continuation for all parabolic subgroups whose unipotent radical is abelian.

The hypothesis has been proved by Gomez and Wallach [2012] for the case when the stabilizer of the unipotent character is compact, and proved by Gomez [$\geq$ 2017] for the general case. The second paper is still in preparation; this is why we write it as a hypothesis.

Let F be $\mathbb{R}$ or $\mathbb{C}$ and π be a generic representation of $\mathrm{GL}_6(F)$ of the form $\pi=I_{\bar{P}}^{G}(\pi_0)$ for some generic representation π_0 of $M(F)=(\mathrm{GL}_2(F))^3$. By the discussion in Section 5B and 6A, we know that in order to prove Theorem 1.4(2) and Theorem 1.5(1) for π, it is enough to show that

$$m(\pi_0) \neq 0 \Rightarrow m(\pi) \neq 0, \tag{7-1}$$

where $m(\pi_0)$ is the multiplicity for the trilinear GL_2 model.

Let $Q_{4,2} = L_{4,2}U_{4,2}$ be the parabolic subgroup of $\mathrm{GL}_6(F)$ containing $\bar{P}$ of type $(4, 2)$, and let $\pi_1 = I_{\bar{P}\cap L_{4,2}}^{L_{4,2}}(\pi_0)$. Then in order to prove (7-1), it is enough to show

$$m(\pi_0) \neq 0 \Rightarrow m(\pi_1) \neq 0, \quad m(\pi_1) \neq 0 \Rightarrow m(\pi) \neq 0, \tag{7-2}$$

where $m(\pi_1)$ is the multiplicity for the middle model defined in Section 5A. Note that the unipotent radicals of $Q_{4,2}$ and $\bar{P} \cap L_{4,2}$ are all abelian. Therefore by the hypothesis, the generalized Jacquet integrals associated to $Q_{4,2}$ and $\bar{P} \cap L_{4,2}$ have holomorphic continuation. This allows us to apply the open orbit method as in Sections 5 and 6, which give the relations in (7-2). This proves (7-1), and finishes the proof of Theorem 1.4(2) and Theorem 1.5(1) without the assumptions on Q.

Acknowledgement

I would like to thank my advisor Dihua Jiang for suggesting this problem. I would like to thank Omer Offen for helpful discussions on the open orbit method. I would like to thank Nolan Wallach and Raul Gomez for answering my questions on the generalized Jacquet integrals. I would like to thank Wee Teck Gan for pointing out that Conjecture 1.2 can be formulated for general representations with nontrivial central characters.

References

[Bernstein and Zelevinsky 1977] I. N. Bernstein and A. V. Zelevinsky, "Induced representations of reductive p-adic groups, I", *Ann. Sci. École Norm. Sup.* (4) **10**:4 (1977), 441–472. MR Zbl

[Deligne et al. 1984] P. Deligne, D. Kazhdan, and M.-F. Vignéras, "Représentations des algèbres centrales simples p-adiques", pp. 33–117 in *Representations of reductive groups over a local field*, Hermann, Paris, 1984. MR Zbl

[Ginzburg and Rallis 2000] D. Ginzburg and S. Rallis, "The exterior cube L-function for GL(6)", *Compositio Math.* **123**:3 (2000), 243–272. MR Zbl

[Gomez $\geq$ 2017] R. Gomez, "Bessel models for general admissible induced representations: the noncompact stabilizer case", in preparation.

[Gomez and Wallach 2012] R. Gomez and N. Wallach, "Bessel models for general admissible induced representations: the compact stabilizer case", *Selecta Math.* (*N.S.*) **18**:1 (2012), 1–26. MR Zbl

[Jiang 2008] D. Jiang, "Residues of Eisenstein series and related problems", pp. 187–204 in *Eisenstein series and applications*, edited by W. T. Gan et al., Progr. Math. **258**, Birkhäuser, Boston, 2008. MR Zbl

[Jiang et al. 2011] D. Jiang, B. Sun, and C.-B. Zhu, "Uniqueness of Ginzburg–Rallis models: the Archimedean case", *Trans. Amer. Math. Soc.* **363**:5 (2011), 2763–2802. MR Zbl

[Loke 2001] H. Y. Loke, "Trilinear forms of $\mathfrak{gl}_2$", *Pacific J. Math.* **197**:1 (2001), 119–144. MR Zbl

[Nien 2006] C. Nien, *Models of representations of general linear groups over p-adic fields*, Ph.D. thesis, University of Minnesota, 2006, available at http://search.proquest.com/docview/305314280.

[Prasad 1990] D. Prasad, "Trilinear forms for representations of GL(2) and local ϵ-factors", *Compositio Math.* **75**:1 (1990), 1–46. MR Zbl

[Waldspurger 2003] J.-L. Waldspurger, "La formule de Plancherel pour les groupes p-adiques (d'après Harish-Chandra)", *J. Inst. Math. Jussieu* **2**:2 (2003), 235–333. MR

[Wan 2016a] C. Wan, "A local relative trace formula for the Ginzburg–Rallis model: the geometric side", 2016. To appear in *Memoirs Amer. Math. Soc.* arXiv

[Wan 2016b] C. Wan, "Multiplicity one theorem for the Ginzburg–Rallis model: the tempered case", submitted, 2016. arXiv

[Wan 2017] C. Wan, *A local trace formula and the multiplicity one theorem for the Ginzburg–Rallis model*, Ph.D. thesis, University of Minnesota, 2017, available at http://www-users.math.umn.edu/~wanxx123/docs/Thesis.pdf.

Received September 20, 2016. Revised March 1, 2017.

CHEN WAN
SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455
UNITED STATES

wanxx123@umn.edu

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

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.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 291 No. 1 November 2017

0030-8730(201711)291:1;1-Y